ECONOMIC ANALYSIS

FOURTH EDITION

VOLUME II

MACROECONOMICS

KENNETH E. BOULDING

PROFESSOR OF ECONOMICS, UNIVERSITY OF MICHIGAN

jointly published by
HARPER & ROW, New York, Evanston & London
and JOHN WEATHERHILL, INC., Tokyo

This is a new and enlarged two-volume edition of a work which originally appeared in 1941 in both a single-volume and a two-volume edition and was later revised in single-volume form in 1948 and 1955. Copyright 1941, 1948, 1955 by Harper & Brothers.

This Harper International Student Reprint is an unabridged, photo-offset reproduction of the American edition. It is for sale in the United Kingdom and on the continent of Europe by HARPER & ROW, LTD., 69 Great Russell Street, London W.C. 1; in Australia and New Zealand by PUBLISHERS SERVICE CENTER, PTY. LTD., 31 Albany Street, Crows Nest, N.S.W., Australia; and in other designated countries by HARPER & ROW, PUBLISHERS, INCORPORATED, 49 East 33rd Street, New York, N.Y. 10016 / Library of Congress catalog card number 66-10054
Printed in Japan / First reprint edition, January, 1966.

TO THE MEMORY OF

Elizabeth Ann Boulding

(1880–1961)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE TO THE FIRST EDITION

The purpose of this book is twofold. It is intended as a text from which the student can learn and the teacher can teach the methods and results of economic analysis. It also seeks to be a contribution to the development and systematization of the body of economic analysis itself. These purposes are not separate. The task of presenting a systematic, orderly, and accurate account of economic analysis is identical with the task of preparing the material for teaching. It must be emphasized, however, that the purpose of this work is not primarily to entertain the student, or to enable him to regurgitate appropriate material into examination books, or to learn a few pat phrases, or to indoctrinate him with an abstract discipline which he will never use. Economics is like photography in this respect, that underexposure is less desirable than no exposure at all; and it is to be feared that too many half-exposed students are produced by our institutions of higher learning. The picture of economics held in their minds is a blurred and confused one, and what they have learned is not sufficiently accurate to serve as a tool for the analysis of practical problems. It is hoped that the student who survives this book will at least have come to regard economic analysis as a discipline useful in the interpretation and solution of numerous problems of life and thought, and will be able to add its methods to the cutting tools of his mind. Hence this is a work for the serious student, and not for the course-taster.

Economics presents a peculiar problem of exposition in that its various parts are much more closely related than is the case in many studies. Unless the student catches some vision of the whole great globe of analysis, therefore, he is likely to miss the significance of each part that he studies. It is the experience of most economists that on first approaching the subject it seems to be a hopeless confusion of unrelated principles. After a certain period of study, however, it may happen that the student experiences an illumination, often quite sudden, of the true nature and relationships of the subject, and from that time forth every part that he studies falls into its proper place and the subject is seen as a closely integrated whole. For this reason I have divided the work into two parts,

each of which contains about enough material for a half-year's course of study. In the first part I have endeavored to range over the whole field of economic analysis, using the simplest possible weapons of analysis, so that the student may obtain a rough outline of the whole picture. In this part, therefore, I treat the concepts of demand and supply curves as self-evident, without discussing the marginal analysis which underlies them. With the concepts of demand and supply as the principal instruments of analysis, I outline the main principles of price determination and of distribution. In this part also I consider the theory of money, banking, international trade, and the business cycle, still confining myself to the tools of demand and supply.

In the second part I give in detail the marginal analysis underlying the demand and supply curves, including the theory of the individual firm, of consumption, of imperfect competition and monopoly, together with a discussion of the theory of capital. With the foundation laid in Part I the student will be able to integrate this more difficult material into his studies as he goes along. In this part I have followed the practice of segregating the more difficult material into separate chapters. Hence this part can be studied on two levels: the student can go through it on an elementary level, or he can take the more advanced material in his stride.

The arrangement of this work thus is not according to subject matter, as usually seems to be the case, but according to the methods of analysis used. The old fourfold division—production, consumption, distribution, and exchange—has almost completely disappeared. Instead we have a twofold division: into the part of the analysis which can be conducted with the aid of the demand and supply concepts, and the part that requires the concepts of the marginal analysis. A surprising amount can be accomplished with the aid of demand and supply analysis alone. Determination of prices, the elementary theory of distribution and exchange, the elementary theory of money, international trade, and the business cycle can be discussed without once using the word "marginal." Therefore, this word so full of torment for the beginner does not occur in Part I. In Part II all this analysis can be elaborated; demand curves can be derived from utility or from production functions, supply curves from cost curves, and the whole theory of the individual firm can be worked out. A possible defect in this method is that the theory of monopoly and imperfect competition must be postponed to Part II. Nevertheless, there are sound reasons for this; a student who is introduced to the modern theory of the market and the individual firm before he receives a thorough grounding in supply and demand analysis is often inclined to lose sight of the broad principles governing the whole economy through his absorption with the study of special cases. The conclusions of the analysis of

perfect competition are modified, but not superseded, by the introduction of more complex and realistic assumptions; and the student will perceive more clearly the implications of these modifications if he knows what they modify.

The method of this work, therefore, is somewhat new; it may be called the "implemental" method, as it seeks to classify the various topics of analysis according to the analytical tools or implements used. It is thus hoped that the student will come to regard these tools as instruments, not as playthings. To this end I have introduced a system of discussion of practical problems which has, I believe, some claim to novelty. Instead of having one volume on "principles" and another volume on "problems," as is usually the case, and where all too frequently the "problems" have little or nothing to do with the "principles," I have integrated principles and problems throughout. That is, after each theoretical section in which the student has been introduced to a tool of analysis, there will be found a section dealing with those practical problems to which the tools thus far acquired can be applied. For example, after the chapter on the determination of prices in a competitive market, I discuss immediately the practical problems of organized competitive markets, such as the foreign exchange market, the capital market, and the commodity markets. After the section on normal demand and supply curves I have two chapters of problems which may be solved with the aid of these concepts, and so through the book.

In selecting these problems I have sought to choose those which provide the best illustrations of the principles involved, rather than those whose interest centers merely in their topicality. It is my belief that a work on principles should compete with neither the popular magazines nor the encyclopedias. Consequently I have not endeavored to write a compendium to current economic problems, for the reason that by the time the student has to face economic problems those of today may no longer be current. It seems to me more important to give the student a training which will fit him to understand the problems of the world of his maturity rather than of his youth. Consequently it is more important to give him a rigorous training in methods of analysis than to prime him with personable current opinions. It also seems to me that it is unwise to crowd a principles course with masses of factual material in special studies —labor, marketing, etc.—merely for the sake of giving the work an air of factuality. The place for such factual studies is later in the student's career, when he has acquired the techniques for interpreting the monstrous riddle of the factual material.

No advanced mathematical knowledge is required for this work. A knowledge of plane geometry will carry the student through all except

the advanced chapters in Part II. Even in these chapters a nodding acquaintance with solid figures is all that is required. Any material involving algebra or the calculus is carefully segregated in appendices.

K. E. B.

Hamilton, New York
March, 1941

PREFACE TO THE FOURTH EDITION

It is now ten years since the appearance of *Economic Analysis,* Third Edition. It is not surprising, therefore, in the light of what has happened during this interval, that the Fourth Edition should require substantial rewriting and rearrangement of topics. The Fourth Edition has been divided into two volumes subtitled Volume I: *Microeconomics* and Volume II: *Macroeconomics.*

Fewer changes from the Third Edition were needed in Volume II: *Macroeconomics.* This perhaps reflects a certain stability in this part of the field, which has not shown any great advances in theoretical construction in the last ten years, although a great deal has been learned in the way of application of econometric models and computer technology that is beyond the scope of this book. Nevertheless, there has been extensive rewriting: the materials have been brought up to date; the treatment of national income and the national accounts has been expanded; and a chapter on Positive and Normative Economics has been added which incorporates some material from Chapter 33 of the Third Edition.

Volumes I and II can be read independently of each other, but it would be preferable to read Volume I first. Although Volume II can be read without the first volume as a prerequisite, it contains a few passages which cannot be fully understood without reference to the volume on Microeconomics. For example, to gain a deeper understanding of the theory of the bank and of financial institutions presented in Volume II, it is desirable to know something about the theory of economic behavior, especially as set forth in Chapters 15 and 16 of Volume I. Most of Volume II, however, can be understood without a detailed knowledge of the intricacies of the marginal analysis.

It is a grave responsibility to write a textbook, and as I meet students from all around the world who have read earlier editions in many different places and many different languages, I am sometimes almost frightened by the thought of how my errors might be multiplied. Many economists have told me that they first "cut their teeth" on *Economic Analysis.* I hope it can still be used as a teething ring, and I hope nobody will ever

mistake it for milk. A book like this is not intended to be swallowed. It is intended rather as an exercise from which the student may emerge with his intellectual tools sharpened, ready to construct his own view of the world.

KENNETH E. BOULDING

Ann Arbor, Michigan
September, 1965

ECONOMIC ANALYSIS

FOURTH EDITION

VOLUME II

MACROECONOMICS

CHAPTER I

ECONOMIC AGGREGATES AND AVERAGES

THE DISTINCTION BETWEEN "MACRO-" AND "MICRO-" ECONOMICS

That part of economics which studies the overall averages an aggregates of the system is often called "macroeconomics." This is often contrasted with "microeconomics," which studies particular firms, prices, outputs, incomes, and expenditures. The distinction, however, is more one of method than it is of subject matter, and, indeed, the subject matter is frequently parceled out somewhat arbitrarily between the two divisions. Thus we might call "microeconomics proper" the study of particular economic organizations, their behavior and immediate interactions with their environment, plus the study of the forces which determine particular relative prices. This leads, however, into "general economics" —the study of the general equilibrium and dynamics of the whole economy. Macroeconomics proper is a particular method for the study of general economics through the use of models involving large aggregates and averages of economic quantities. Traditionally, however, the study of money and finance, including the theory of the behavior of banks and other financial institutions, has been included in macroeconomics, perhaps because money and finance is such a crucial part of the general economic system, even though to be consistent we should include the theory of the financial institution in microeconomics, for it is part of the general theory of economic organizations just like the theory of the trading or manufacturing firms. We shall follow tradition in this respect, obeying convenience more than a strict consistency; nevertheless we shall find that the theory of the financial institution

does not differ essentially from the general theory of the economic organization (See Volume I: *Microeconomics,* Chapter 15) and can easily be seen as a special case of the more general theory.

The Place of the Economics of Government

Similarly the study of governmental organizations of all kinds as economic organizations strictly should fall under microeconomics, for again governmental organizations, whether municipalities, states, national governments, or government corporations like the TVA are special cases of the general species of economic organizations and follow some of the principles of the theory of economic behavior. Nevertheless the economics of governmental organizations usually is studied under macroeconomics. There is some logic to this, for a government which is not arbitrary or tyrannical must deal with individuals not *as* individuals but as members of some group or aggregate of individuals; this is what is meant by the "rule of law." It can be said also that the main economic responsibilities of government lie in the regulation of the aggregates of the system—general prices, general outputs, the general volume of trade, and so on—and that the more it intrudes itself into the detailed regulation of particular prices, wages, and the like, the less successful is its policy likely to be.

"Slicing" and "Lumping"

Even though macroeconomics and general economics both study the total economic system, they approach this problem by different methods. General economics studies first small pieces or sections of the system, with a few isolated variables, and then attempts to generalize from these subsystems to the system as a whole, using mathematical inference. This I have sometimes called the method of "slicing"—we take a slice in two or three or a few dimensions of the enormous n-dimensional reality and study it carefully, hoping to learn something of the whole system by the analysis of sections of it. This is rather like the brain physiologist's approach to the problem of the human person. Macroeconomics by contrast uses the method of "lumping"—dividing the whole system into huge chunks and studying the relations among these aggregrates; this is more like the psychoanalytic approach to the human person. Both these methods have their value. The methods of general economics are more elegant, but it is hard to get tangible results from them; the methods of macroeconomics are crude, but yield important insights. Macroeconomics is "bulldozer" economics; it is designed to push large problems around in a crude way, and it is often only weakened by excessive refinement.

Difficulties in Macroeconomic Analysis

There are certain difficulties and dangers in macroeconomic thinking which should command attention. In the first place, many propositions which are true of individuals or of small groups turn out to be untrue when we are considering the system as a whole. There are many things also which an individual can do only because most other individuals refrain from doing it. Most of us, having money in our wallet, feel able to go to a movie or take a train ride. Yet, clearly, if everyone tried to do these things at once, there would be no seats available for most people either at the movies or on the train. Similarly, in normal times anyone who has a bank deposit is free, and able, to go to the bank and exchange it for cash; but if every depositor did this at once, the bank would have to close. Any individual can increase the amount of money he has by simply not spending as much money as he receives. Unless there is actual creation of new money, however, it is impossible for all people to spend less than they receive; for every expenditure is at the same time a receipt to the person to whom it is made. In macroeconomics, therefore, we must be on our guard against generalizing from our individual experience; just because we ourselves can do something is no reason for supposing that everybody can do it at the same time. Generalizing from our own experience is such a common habit that we constantly fall into it; it is, however, one of the greatest sources of error in social thinking.

Dangers in Aggregative Thinking

The second danger in macroeconomic thinking is of a quite different kind; it is that we may think too easily in terms of aggregates as if they were homogeneous, without realizing the significance of their internal composition and structure. We see many examples of this in political thinking. Constantly in discussion we speak of France, Russia, or the United States, as if these aggregates of people were single entities. For some purposes they can, indeed, be so regarded, and unless they are so regarded discussion becomes almost impossible. We cannot, every time we mention the United States, substitute "The people living on North America between Canada and Mexico, their separate habits, characters, divisions, groups, classes, parties, institutions, and organizations." Nevertheless, it is easy in discussion of any kind to forget that shorthand symbols like the United States, in fact, stand for a great, complex diversity of men and institutions, which for some purposes, but not for others, can be regarded as a single entity. In macroeconomic discussion likewise it is easy to forget that the aggregates

or averages under discussion are in fact made up of innumerable individual times, and that changes in their internal structure or composition may be more significant in the interpretation of some particular problem than changes in the aggregate itself. To take but a single example: Many economists have tried to postulate a relationship between the average wage and the volume of employment, on the analogy of the demand for labor of a particular kind. The average wage, however, is derived from a very large number of particular wages—e.g., wages of carpenters, machinists, nurses; and the volume of aggregate employment, in so far as it depends upon wages at all, may depend much more on the relative structure of wages than on the average level. If, for instance, wages of carpenters rose but wages of nurses fell, the average might remain unchanged; but if the employment of carpenters fell only a little, whereas the employment of nurses rose a great deal, aggregate employment would rise.

Aggregates Must Be Interesting

A third danger in aggregative analysis is that the aggregates which compose our system may not be significant or interesting. An aggregate consists of all the items in a given universe which conform to some definition. A definition is a kind of verbal hedge or wall—it sets up an enclosure inside of which stand all the things which conform to the definition and outside of which stand all things which do not so conform. There is no limit to the number of such enclosures which can be made. Some of them, however, are interesting and some are not. Suppose, for instance, we form a concept, say, of the aggregate of all left-handed people who have red hair. We must, of course, find a name for the concept, usually from Latin or Greek. A left-handed person with red hair, then, we call a "sinirufe." There will be some argument of course as to how left-handed and how red-headed a person has to be before he can qualify as a sinirufe, as practically all definitions are a little vague—the enclosure is divided off from the outside world by a fog rather than by a fence. One might even imagine learned disputes among the sinirufists as to what constitutes a "real" sinirufe. Sinirufes, of course, once we agreed on a definition, could be counted, and statistics, even time series, of the number of sinirufes compiled. Studies could be made of their geographical distribution, their heredity, their marital histories, their aggregate and per capita income, their wealth. Doctoral dissertations and textbooks could be written about them, professors of sinirufics established, even interdisciplinary research developed between sinirufists and related disciplines. In fact this does not happen, even though some things which happen in academic circles are only slightly

less ridiculous. It is important to ask why the sinirufe is so neglected, if only to throw light on why certain other definitions, concepts, and fields of study are pursued. The answer is simple: the sinirufe is neglected, indeed was not even named until I named him, because he is not interesting *as such,* however interesting a particular sinirufe might be as a person.

Aggregates Are Interesting Only if They Can Be Functionally Related

Interest is the nutrition of science, and no concept or study will live in the minds of men unless *somebody* finds it interesting. The sinirufe is not interesting in spite of the fact that he exists, and that a whole potential field of study waits to be erected around him. The sinirufe is just as "real" as the plumber, the Seventh-Day Adventist, the American, the Jew, the Negro, the economist. Like these he could be identified, seen, touched, smelt, heard, and perhaps even, in some cultures, tasted. He is not interesting because he has no function. Sinirufes do not do anything, and do not have anything done to them, as such. They are not perceived by others as constituting a group, nor do they perceive themselves as a unity. Their behavior as an aggregate is simply the sum total of the individual behavior of particular sinirufes: it cannot be assumed to depend on the behavior of any other aggregate. In mathematical language, if S is some quantity descriptive of the aggregate of sinirufes—their number, weight, income, I.Q., or any other quantity—it makes no sense to write that S is a function of H $[S = F(H)]$ where H is any other quantity in the universe. If it was discovered that such functions or relationships existed in any society, even if the sinirufe were not perceived as such in the society, the concept would immediately have scientific interest. Suppose for instance that all or most sinirufes were more inventive, or stupider, or more prejudiced than people who are not sinirufes. Immediately the concept would become interesting, for the percentages of sinirufes in any group would be directly related to the inventiveness, stupidity, or prejudice of the whole group. The sinirufe index might become almost as useful, and even more meaningful, than the intelligence quotient.

Difficulties of Definition

I have used what may seem to be an absurd or even frivolous example above because so many grave discussions in economics as in other sciences are almost as absurd. A great deal of breath and ink have been wasted in economics, for instance, on the question of the correct definition of concepts associated with words like capital, income, value, and so on. For a long time it was customary in text-

books of economics to begin with a chapter of definitions and concepts. The definitions given in the first chapter were usually forgotten by the middle of the second chapter, but at least the author had paid homage to the great principle of linguistic monogamy—one meaning wedded to one word. Unfortunately this is a principle quite inadequate to take care of the number of meanings wanting words, and some form of linguistic polygamy, or at least Hollywood marriage, seems to be in order.

Economic Dimensions

About the only distinctions which are quite clear, and about which there can be no argument or shadow of doubt, are those based on differences in *dimensions.* In Newtonian physics, for instance, all quantities can be reduced to some combination of powers of the three basic dimensions: Length (L), mass (M), and time (T). Thus distance is simple length, L, or so many feet. Velocity is measured as a length per unit of time (L/T) or as so many feet per second. Acceleration is the change in velocity per unit of time, and is measured therefore as a length per unit of time per unit of time (L/T^2) or as so many feet per second per second. Distance should never be confused with velocity, or velocity with acceleration; the difference in dimensions between these concepts constitutes a clear and unscalable barrier between them. Similarly in economics there are three basic dimensions: commodity (C), money (M), and time (T). The commodity dimension is in fact multiple: it should really be written $C_1, C_2 \ldots C_n$, each symbol standing for the different physical measure of different commodities—tons of steel, bushels of wheat, bars of soap, units of Chevrolets. The money dimension likewise may be multiple: it may stand for dollars, pounds, francs, etc.

All economic quantities can be expressed in terms of some combination of these dimensions.[1] Thus capital in physical terms has the dimension C: it is always a stock of some physical objects, defined at some instant of time. Income on the other hand has the dimensions C/T. Capital is a number of bushels or tons of something; income is bushels or tons per day or per week or per year. The capital and income concepts are therefore quite distinct, and should never be confused. It is no exaggeration to say that the major source of confusion and error in classical economics is the constant confusion between the capital (stock, fund) concepts and the income (flow) concepts. The wage-fund theory in its crude form exhibits this confusion, as it identifies a stock (a portion of capital) with a flow (the wages bill)—

[1] For some purposes an additional dimension, utility or satisfaction, U, is necessary. Most economic theory, however, can be developed without it.

two things which cannot from their very nature be identified, though they may be related.

Some other dimensions of economic quantities are of interest, and may briefly be noted here. A ratio of exchange or value is the ratio of two quantities of commodity, or C_1/C_2. The special case of a price is a ratio of a quantity of money to a quantity of commodity, or M/C_1. We can, if we like, regard money as simply a special case of commodity, in which case it would be included in the commodity dimension. Its peculiarities, however, perhaps make it worthy of a special symbol. The concept of commodity here includes any asset whatever which can be exchanged; it includes therefore such things as securities, bonds, futures contracts, mortgages, and any other exchangeable financial instrument. The dimensions of the rate of interest are perhaps a little surprising at first sight. The rate of interest is measured as a percent per annum. The "percent" has no dimensions at all, as it is a ratio of two quantities which have the same dimensions. Thus "5 percent" is simply five "dollars" divided by a hundred "dollars," or 1/20. The dimensions of a percent per annum therefore are $1/T$. Thus the reciprocal of the rate of interest is a period of years (the period of purchase): 5 percent per annum is exactly equivalent to 20 years purchase. The dimensions of the rate of interest can also be visualized as the ratio of a constant income to its capital value—say \$5 per year per \$100. This would be symbolized as $M/T/M$, or again $1/T$. The dimensions of the velocity of circulation of money are likewise $1/T$; its reciprocal is the period of turnover—also a length of time (see page 74).

Definition Within a Single Dimension Depends on Its Purposes

I have emphasized the necessity of maintaining clear and precise distinctions between quantities having different dimensions, both in thought and language. The boundary between capital and income is a "natural" boundary, like an ocean which separates two continents. The boundaries *within* the capital or the income concepts, however, are to a large extent artificial, like the boundaries of states. Definitions within the broad continent of "stocks of things" or the other great continent of "flows of things" are to some extent arbitrary, are matters of convenience, imposed on us in the last resort because of the limitations of language. There is no point, therefore, in discussing what is the "right" concept of capital or of income if we mean by the right concept some particular section or area within the broad field of stocks of things or of flows of things. There are a great many useful concepts within each field, and likewise a great many useless ones. The usefulness of a concept depends mainly on the extent to which

the aggregate defined by it can be related to others, as we have seen. The problem is in many ways analogous to the problem of the geographer in defining a region. Everyone would agree that the "Middle West" is a useful concept in geography, even though it might be difficult or even impossible to reach agreement as to where its exact boundaries lie. Similarly, within the broad field of stocks of goods we might define a region of, say, circulating capital, or goods in process, even though it would be impossible to draw the boundaries of this region in a way that would please everybody.

Overlapping Regions

Just as the regions of a geographer may overlap, depending on the purpose for which they are constructed, so may the concepts of the economist. Thus for some purposes Michigan is a significant geographical entity, because it shares a common state government and state law. For other purposes it is useful to run a boundary right through Michigan, joining its southern part to the corn belt region and its northern part to the cutover forest region. Similarly within the broad field of capital, or stocks of goods, we may define overlapping regions. Thus for some purposes a distinction between household goods and business goods may be valuable, in so far as businesses and households behave differently. Cutting across this distinction altogether is a distinction between durable and nondurable goods. Households possess both durable goods like refrigerators and nondurable goods like sugar. Businesses likewise possess durable goods (factories and machines) and nondurable goods (machine oil, gasoline). The essentially arbitrary nature of these distinctions, and also the fact that their arbitrary nature in no way detracts from their usefulness, is admirably illustrated by this last distinction. It is obviously impossible to draw any clear line between durable and nondurable goods. There is a continuous, or nearly continuous, scale of durability of goods, and any line that we draw is arbitrary. For some purposes we may want to draw several lines —for instance, we may wish to divide the stock of goods into durable, semidurable, and nondurable. The U.S. Department of Commerce does this. We draw these lines, however, mainly because the continuum of the durable-nondurable scale is too complex to talk about. No matter where we draw the line between durable and nondurable, within reason, the study of the changes in the composition of the stock of goods from the standpoint of durability will probably be reflected in changes in the proportion of nondurable to durable goods according to our arbitrary definition.

Dangers in Arbitrary Definitions

The above example will also serve to illustrate the dangers as well as the uses of arbitrary definition. Suppose for instance that we draw the line between durable and nondurable goods at a length of useful life of one year, so that all goods with a length of life of less than a year are classified as nondurable and all goods with a length of life of a year or more are classified as durable. Suppose now that there is a large increase in goods lasting 13 months and a large decline in goods lasting 11 months. This change in the composition of the total stock of goods will be reflected in a sharp rise in the quantity of durable goods according to our definition, and a sharp fall in the quantity of nondurable goods. It will look as if there has been an important change in the durability of the stock of goods in the direction of a rise in durability. In fact the change may be quite insignificant; it is merely the accident that the change occurred across the arbitrary boundary which made it look significant. The measurement of changes therefore at arbitrary boundaries is always open to some question. Changes across these definitional boundaries are accurate measures of the total changes in the composition of an aggregate if the actual changes are uniformly distributed through the total field. Thus if all items in the capital stock increased in durability, then no matter where we drew the line between durable and nondurable goods there would be a rise in the proportion of durable goods as defined, which would have some significance as a measure of the general rise in durability.

Classifications of Goods: Instruments and Goods in Process

Many other classifications of the total stock of goods are possible, based on more or less continuous arrangements of properties of these goods. Thus we may make a distinction between instruments and goods in process. Instruments are those goods which are complete in themselves and which render a succession of ~ervices over a period of time. Houses, factories, machines, automobiles, domestic animals are all examples. These things are frequently called "fixed capital," not so much because they are fixed in position as because their physical form does not change significantly in the course of their service. Such things depreciate in value as they are used, even apart from any minor changes in their physical form. Even the celebrated one-hoss shay, whose physical form remained unchanged for 100 years, would have

depreciated in value between the time of its creation and the dramatic moment of its dissolution.

Goods in process are those things which are not complete in themselves and which must undergo a complete transformation of form before they render services. Raw materials of all kinds, such as wheat, flour, coal, or iron, are good examples of this category. These are frequently called "circulating capital," as they must be completely transformed physically before they yield services. Wheat must be turned into flour, flour into bread, before they can yield satisfaction. The distinction between fixed and circulating capital, like other classifications of capital, is not absolutely clear, nor is it important that it should be so. Taken over 100 miles, the gasoline in the tank of a car is clearly circulating capital and the tires are fixed capital, for although they wear a little their form is essentially unchanged. Over 20,000 miles, however, the tires are as much circulating capital as the gasoline, for they have yielded up their form entirely. The tires as well as the gasoline have been transformed into "miles."

Original Goods and Produced Goods

Another distinction frequently made is that which involves the division of the total stock of goods into original goods and produced goods. Original goods are those which have value but which have not been produced by man. Land is usually regarded as the main example. Minerals, coal, and other minable deposits may also be included. Produced goods, on the other hand, include such things as buildings, raw materials in storage, and equipment of all kinds, which have been created by the activity of man. This distinction is perhaps the most difficult of all to maintain clearly, yet it has played an important part in economic thought. According to Ricardo, for instance, "land" is the "original and indestructible powers of the soil." But in any given case it is difficult, and often impossible, to distinguish between what is original and indestructible and what is produced and mortal. It may be questioned, in view of what neglect and erosion can do, whether there are any properties of the soil which are indestructible, except perhaps in the ever-fertile river bottoms. It may be questioned also whether land is indeed not "produced." Is not the labor and investment of the pioneer whose sole reward is the ownership of land an act of "production" of land? It will not do to assume too hastily that there is a vital and clear distinction between land and produced goods.

Nevertheless, the distinction has some foundation in reality. There is a difference between those goods which come into being as a result

of conscious effort and those goods which are "windfalls," which are discovered accidentally. The prospector who finds a gold mine after twenty years of heartbreaking search has perhaps produced the gold mine as a result of his activity. The man who finds a gold mine by accident in his back yard has not produced it, for it has not come into being as a result of purposeful activity. The importance of this distinction rests in the fact that the supply of genuinely original goods is perfectly inelastic. The quantity of them which will be forthcoming depends in no way on their price. The probability of discovering the gold mine in the back yard, for instance, depends not at all on the price of gold. But the supply of produced goods will always have some degree of elasticity, over a sufficiently long period of time. If the price of gold is high, presumably more prospectors will be encouraged to pursue their labors, and therefore the probability of discovering gold is greater than it would be if the price were low.

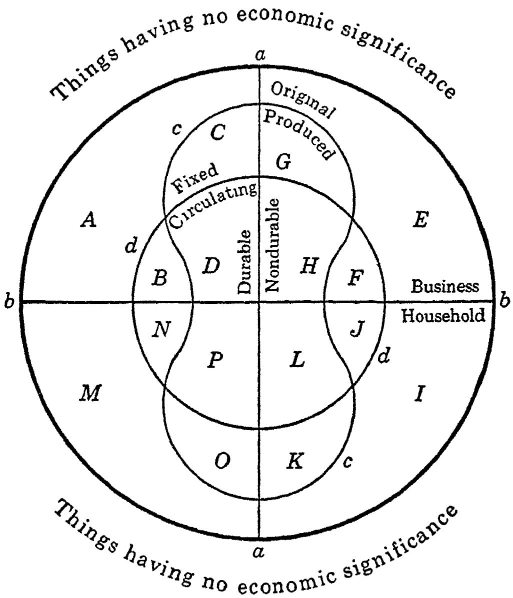


Fig. 1. Classification of Goods

A Classification Diagram

It will perhaps be helpful if these distinctions are shown in diagrammatic form, as in Fig. 1. We suppose first that all things in existence at a moment of time are spread out in the "field" of the paper. The

thick black line then represents the definition of capital in the broadest sense as the total stock of things at a given moment having economic significance. The shape of the figure is of no importance; I have drawn it as a closed figure, however, to show that the number of things having economic significance is limited, whether the total universe is infinite or not. It is not always perfectly clear exactly where this boundary lies, but there will be agreement as to its general location. Things which are priced, are actually exchanged, or are offered for exchange are clearly within the boundary. Things which are potentially priceable or exchangeable probably should be in it. Things which may influence things which are exchangeable are more doubtfully included or excluded. Argument about the exact position of the boundary, however, does not invalidate the concept of a region. Within the region, therefore, we can distinguish subregions by drawing further boundaries. In the figure I have drawn four such subboundaries, each dividing the total region into two parts. The line *aa* separates the region into durable and nondurable goods, *bb* into business and household goods, *cc* into original and produced goods, *dd* into instruments (fixed) and goods in process (circulating). Clearly many other divisions could be made. With four such divisions the maximum number of classifications or "cells" in the figure would be 16, or 2^n, where n is the number of twofold divisions. Thus all goods in area A are durable, business, fixed, and original goods. The size of the various "boxes" is not significant in this figure, and some of the boxes will actually be empty. Classification systems do not have to be based on dichotomies like the above; we can have any number of classes of each dimension. We can even have an infinite number of classes of each dimension, in which case it becomes a continuous variable. Thus in Fig. 1 we might abolish the line *aa* and simply say that the farther to the left any point is, the more durable the goods.

THE INCOME CONCEPTS

The problems of classification and definition which we have examined briefly in the case of the capital concepts are present to an even greater extent in the case of the income concepts. The general distinction between the capital and income concepts is, of course, a very clear one, based as it is on a difference in dimensions. Any quantity which is measured without reference to a period or length of time, but only with reference to a date or instant of time, is a "stock" or capital quantity. Any quantity which is measured as so many units *per unit of time*—dollars per year, bushels per day, gallons per second—is a flow or income quantity.

Production, Consumption, Accumulation

Corresponding to the broad concept of physical capital as the total inventory of economically significant objects we have three flow concepts—production, consumption, and accumulation. The total inventory of stock may be compared to a gigantic "pool" of things. Into this pool new things are constantly pouring from the productive process. Bushels of wheat are added to it as they are grown, piglings are added to it as they are farrowed, pigs are added to it as piglings grow up, buildings are added to it as they are built, and so on. Out of this pool things are constantly draining away into consumption. Food is eaten; clothes are worn out; coal is burned up; machines grow obsolete, depreciate, and are scrapped. Production, therefore, is the process of adding to the total inventory of stock. Consumption is the process of subtracting from this total inventory. The difference between the rate of production and the rate of consumption may be called the rate of real accumulation; it is the rate at which the stock is increasing. If wheat is being produced at a rate of 100 million bushels a year and consumed at the rate of 90 million bushels a year, then the stocks of wheat are increasing at a rate of 10 million bushels a year.

The Problem of "Netness"

In the case of a single commodity there is not much difficulty in identifying production as the gross rate of addition to the total stock of the commodity, consumption as the gross subtraction from the total stock, and accumulation as the net additions. Even in this case, however, a certain difficulty may arise where some of the stock is used in the production of further stock. Thus scrap iron is used in the production of further iron. Should the consumption of scrap iron in the furnaces be counted in the total of consumption and the iron produced from these same furnaces be counted as production? Or should the iron used in the furnaces be deducted from the iron produced by the furnaces to obtain a figure for *net* production? This problem of what constitutes "netness" in production is one of the most troublesome in the theory of income. If the definitions are consistent the measure of accumulation is not affected, but the more "gross" our concept of production the larger will be the totals of both production and consumption. Thus in the case of iron, suppose that 100 million tons are produced at the furnaces, 25 million tons of scrap are consumed in the furnaces, and 10 million tons are lost or consumed by rust. We can, if we like, say that 100 million tons have been produced and 35 consumed, with a consequent addition of 65 to the total stock of iron; or we can say that net production

was 75 million tons, net consumption 10 million, again with an accumulation of 65 million tons.

Depreciation and "User Cost"

When we consider what should be included in the aggregate of all production and consumption, the problem of how "gross" our measures should be becomes an acute one. Wheat is made into flour, flour into bread. Should we count the wheat and the flour and the bread, or should we count only the bread in the total of production, on the grounds that to include the wheat and the flour would result in double counting, since the wheat is simply transformed into the flour and the flour into the bread? We may feel fairly sure of the answer in the above case, as the "net" concept here seems to be clearly more interesting and significant than the "gross" concept. The problem does not end at this point, however. Some machinery will have depreciated in the course of producing the flour and the bread. This depreciation is clearly part of total consumption, yet should it be deducted from the production of bread in order to arrive at an aggregate of net production? It would seem that if we deduct the flour we should also deduct the wear and tear on the machinery, for both are equally "real costs" of producing the bread—both represent a diminution in the stock of flour or machinery in order to increase the stock of bread. There is, however, a possible difference. If the bread were not made, the flour would presumably remain intact, short of damage in storage. If the bread were not made and the bakery stood idle, however, it would still depreciate, perhaps less —or perhaps even more—than it would if it were used, through the inexorable processes of physical and chemical decay, the "iron law" of moth and rust. Perhaps therefore we should deduct only that part of depreciation which can be clearly attributed to the production (called by Keynes "user cost") in defining our net product.

Subsistence as a Cost

The difficulty, however, does not even rest at this point. Production almost inevitably involves labor, and labor cannot be performed unless the worker is fed and clothed. Should, therefore, the subsistence of the worker be deducted from the total of production in order to arrive at a net product? Ricardo and the other classical economists certainly thought so. We do not count the food of the cow in the total net product, only the milk which that food produced by being, as it were, passed through the cow. Why, then, should we not also deduct the food of the stockman and the milker from the gross product, as this equally goes to produce the energy which is just as necessary as

the feed in the production of the final product? At this point, however, the modern economist generally revolts: in a democratic society we cannot lump together the "laboring poor and the laboring cattle" as Adam Smith did, and regard the subsistence of both as simply a real cost which must be deducted from the final product in order to arrive at the true net product. Sentiment apart, however, there is much to be said in logic for the classical point of view.

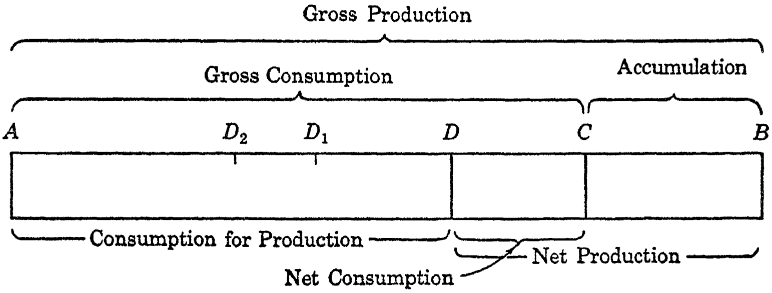


Fig. 2. Gross and Net Product

Accumulation Concept Not Affected by "Netness" of Income

The above concepts are illustrated in Fig. 2. Suppose that the length of the line AB measures the gross production of a commodity, say flour—that is, the total number of units of flour created during a given period of time, say a year. AC represents the gross consumption of flour—that is, the total number of units of flour destroyed, by any method and for any purpose, during the same period of time. CB then represents the accumulation of flour, or the net increase in the stock of flour during the same period. If consumption exceeds production, CB will, of course, be negative, indicating a decumulation or diminution in the total stock of flour. Now we suppose that a certain amount of the consumption of flour, AD, is consumed in the production of other things (or of flour itself). DB is then the net production of flour, DC the net consumption. CB still represents the amount of flour accumulated. It is clear that CB is equal to the difference between net production and net consumption, no matter where we place the point D. For different purposes we may wish to place the point D at different levels, say D_1, D_2, etc.

Goods of Zero Length of Life

If now we wish to get a picture of the aggregate production-consumption-accumulation patterns in real terms, we need to visualize a figure like Fig. 2 for *every* commodity. For every commodity we can postulate

a gross addition to stocks (gross production), a gross subtraction (gross consumption), and various definitions of net production and consumption, depending on our definition of "consumption for further production." We may find some items which seem at first sight difficult to fit into this picture, but we shall find nothing in economic life which is basically inconsistent with it. Thus there are some commodities (services) which have so short a length of life that we can hardly regard them as having any "stocks" at all. Production and consumption in this case occur simultaneously, and it seems a little odd to regard production as the act of adding to zero stocks, an addition which is immediately subtracted in consumption. Nevertheless, by so regarding the production of services, we can fit the concept into the general pattern without any distortion of reality.

Services and "Psychic Capital"

From one point of view, indeed, it can be argued that all things which we regard as services involve the creation of capital. Thus a haircut involves the creation of a capital good—the barbered head—which inexorably depreciates as the days go by and must be replaced after a period suitable to the profession of the subject. The growth of a discommodity, whether it is unwanted hair or unwanted grass, is as much depreciation as the decline of a commodity. Even in the case of services where the embodiment is not apparent to the senses we can suppose that "psychic capital" is created which also depreciates and eventually has to be replaced by another act of production. Thus attendance at an opera produces a "psychic good" which might be called "just having heard an opera." As long as this psychological state is in good repair we feel no urge to go to another opera. The state, however, decays, and after a period we find ourselves in the state of "not having been to the opera for a long time," whereupon we go to hear another opera and restore the original state of agreeable musical fullness. It is pleasant to speculate whether the memory of a bad opera lasts longer than that of a good one; however that may be, it is clear that the rate of depreciation varies greatly from person to person. For the ardent fan the state of "just having been to the opera" decays overnight; in another the state may be kept in repair quite adequately by an annual visit.

The "Product" of Government

Some difficult problems are presented in the field of income concepts by government activity, mainly because the "product" of government activity in real terms is difficult to measure, and cannot be measured directly, and also because it is difficult to say how the product of govern-

ment should be apportioned among the various beneficiaries. At one extreme the cost of government might be regarded as a sort of "overhead" of society at large, providing something of the framework within which the productive process is carried on, but not in itself constituting a product. On this assumption the whole cost of government would be deducted from the gross product in order to calculate a net product. At the other extreme one might assume that government produced an intangible but real product which was apportioned among the people in proportion to the taxes paid. This assumes that corresponding to the flow of taxes to government, there is a flow of some invisible real product for which the taxes are payment. Neither of these extreme assumptions is at all satisfactory, and in the practice of national income accounting uneasy compromises must be made between them.

Product and Welfare

Neither individual nor aggregate welfare is susceptible to direct measure. Nevertheless, it is reasonable to suppose that there are certain positive correlations of a rough kind between various measures of aggregate product or consumption and the overall welfare of a people. As between a country with a per capita income of $100 and one with $2000, we are fairly confident that the latter is better off than the former. Nevertheless, no measure of aggregate product is an exact measure of welfare, and we have to avoid certain pitfalls in using it as such. Welfare is a "state" of the person or aggregate of persons; it is a stock concept rather than a flow concept. We assume that production or consumption are measures of welfare only because of a prior assumption that the more extended and elaborate the state of welfare the more consumption it involves and the more production is needed to maintain it. Any improvement in the durability of capital, whether physical or psychic, therefore, will in itself increase welfare, and we will get more welfare per unit of production or consumption. Furthermore, consumption is rarely a good in itself; it is a "cost" of maintaining certain states, and production is what has to be done to overcome this cost. Measures of production or consumption, therefore, can easily fail to measure welfare on this account. Thus cold countries have to produce and consume fuel for space heating just to keep warm, and because of this perhaps have a larger product than warm countries which do not have this necessity; this, however, is a sign of poverty, not of riches! Measures of product likewise fail to take into account the varying demand for leisure; a country with an agreeable climate and lovely scenery may have a lower product simply because its people prefer to produce less and enjoy themselves more. This is perhaps part of a more general problem in which the existence of free goods,

or goods which have a zero price in some cases but not in others, seriously distorts the meaning of the aggregate product, for free goods do not appear in it and yet they contribute to welfare.

AGGREGATES OF CIRCULATION AND EXCHANGE

The concepts of production, consumption, and accumulation are flow or income concepts in the sense that they are measured in units per unit of time. There are, however, other important flow concepts in economics which correspond more closely perhaps to the physical analogy of a stream. These are the flows of circulation or exchange, and also of geographical movement. So many "bushels of wheat per year" might refer not only to quantities produced, consumed, or accumulated; it might also refer to quantities flowing across a boundary, or between two points, or between two owners or sets of owners, or from farms to warehouses. Literal streams of wheat flow from the wheat fields to the cities; streams of automobiles flow from Detroit to the ends of the earth. These flows constitute the total volume of physical trade. Against them flow corresponding "streams" of means of payment (money). Even though the money flows may not correspond geographically to the flows of commodities, they always consist of transfers of money per unit of time between economic organisms, opposite in direction to the flows of the things for which money is paid.

Distinction Between Income and Exchange Concepts

Although they have the same economic dimensions, there is a clear and vital distinction between the production, consumption, and accumulation concepts on the one hand and the trade and money flows concepts on the other, and much confusion has been caused in economics by a failure to keep this distinction in mind. Production and consumption measure the creation and destruction of things; trade and exchange measure the extent to which existing things are shuffled around among owners or among regions. In exchange no physical asset is either produced or consumed, but assets shift owners. A *transfer* is a simple shift of ownership from one economic organism to another, such as a gift. An exchange is two transfers in opposite directions between the same two parties. Likewise a money transfer, or *payment,* involves no creation or destruction of money, but simply a shift in its ownership. Every payment, therefore, is both an expenditure and a receipt—an expenditure to the account out of which money is transferred, and at the same time a receipt to the account into which the money is transferred.

Confusion arises among these concepts because a money payment in purchase of a household good measures three things: the amount of the money payment itself, the *value* of the goods for which the money was paid, and the value of the consumption of these goods when they are consumed. For purposes of analysis, however, these three things must be kept distinct. Thus even when somebody buys and eats a 10 cent ice cream three things have happened: (1) a transfer of money (10 cents) from the buyer to the seller, (2) a transfer of 10 cents' worth of ice cream from the seller to the buyer, (3) the consumption of 10 cents' worth of his assets (ice cream) by the buyer. The first two items constitute an exchange. If they are included in aggregates they will be in aggregates of trade or of money flows. The last item is part of consumption, and will be included in the aggregate of consumption.

Exactly similar considerations apply to the important distinction between the *earning* of income and the *receipt* of money. The receipt of a paycheck is not the earning of income, but is again an exchange, though not such an obvious one as the exchanges involved in the purchase of consumer goods. We go to the pay office possessed of a claim on our employer for, say, $100. We return with a check which is essentially a claim on a bank for $100. In the earning and receipt of wages, therefore, three distinct items or operations are again present. The earning of income is the accumulation of a claim against the employer as labor is performed. When wages are paid, this wage claim is given up to the employer. Even though nothing physical changes hands it is generally understood that the payment of wages constitutes a relinquishment of the wage claim. Then in return for the relinquishment of the wage claim the employer pays the worker an equivalent sum in money. These distinctions may seem almost trivial; their importance, however, especially when the individual items are aggregated, will become clearer in later chapters.

Money flows are also involved in payments for securities, stocks or bonds, for money of other countries (foreign exchange), for different kinds of money, such as bank deposits, which are exchanged for other kinds, and so on. All this is known as the "financial circulation."

Aggregates of Prices and Valuation Coefficients

Just as there is a problem of the definition and classification of aggregates of stocks or of flows of goods, so there is also a problem in defining and classifying aggregates or averages of prices. A price is the ratio of a number of monetary units (e.g., dollars) to a related quantity of some physical good. There are two kinds of price ratios which are significant in economics. The most obvious is the ratio of exchange—i.e., the ratio of

the amount of money paid for something in an actual transaction to the amount of goods or other assets paid for by the money. Thus if somebody buys 20 tons of wheat for $40, the price of wheat in the transaction is $40/20 tons or $2 per ton. Prices as quoted and recorded in statistics frequently represent not ratios of exchange in actual exchanges but offers to sell or to buy at the ratios quoted. Thus a price tag in a store is equivalent to an offer to exchange the goods in question at the ratio quoted. Price ratios also exist, however, which are neither descriptive of actual exchanges nor of offers to trade. These ratios may be described as *valuation coefficients.* Valuation is the process whereby we calculate the equivalent in some measure of value of a physical quantity of the valued objects. The valuation coefficient then is the ratio between the value of something, expressed in units of the measure of value—dollars again are an example—and the quantity, expressed in some physical terms, of the thing which is valued. Thus in compiling a list of personal assets we might value two carpets at $400. The $400 is simply a figure expressing the equivalent of the carpets in the measure of value. It does not represent an offer to sell, or a ratio of actual exchange. The valuation coefficient would be $200 per carpet.

It should be observed that exchange ratios, actual or potential, always refer to actual or potential *flows;* valuation coefficients generally refer to *stocks.* In general, valuation coefficients are based in some way or another on exchange ratios, past, present, or future. Thus a valuation may be made by recording the original exchange ratio when the valued article last participated in an exchange (was bought) and making some allowance for depreciation (or even appreciation) since that exchange. Or it may be made by observing the actual exchange ratios of similar objects which are being exchanged at the moment. Or it may be made by noting the ratio at which the object will be exchanged at a future date (say, in the case of a bond) and making some allowance (discounting) for the futurity of the exchange. In any case it is clear that no basis would be present for valuation unless there was in continuing existence a system of actual exchanges. This gives the price aggregates of actual exchanges a peculiar importance.

Price Sets as Aggregates

The price set is the set of all prices as they exist in exchanges or offers to exchange at a moment or during a brief period of time. We have a similar concept of the set of all valuation coefficients. For some purposes we want to divide these sets into subsets, just as we divide the "quantity set" of quantities (whether stocks or flows) of commodities into subsets of different kinds of goods. Thus we may divide the price set into the

set of retail prices, wholesale prices, prices of producer's goods, prices of consumer's goods, and so on. There will be a price set corresponding to any classification we wish to make within the quantity set.

The problem of price levels is that of finding some number or index which is derived from a price set or subset in such a way that it represents the "general level" of prices in the set. This is an average rather than an aggregate, simply because a price is a ratio, not a simple quantity. Conceptually, therefore, a price level is also a ratio of money paid to the quantity of things bought with it. We shall develop this concept in the next chapter.

QUESTIONS AND EXERCISES

1. Make a list of all 16 possible types of goods in the boxes A,B—P of Fig.1. Try to find examples for each category; where examples cannot be found, label the category as "empty." Now redraw Fig. 1 so that there are no empty boxes.
2. Draw figures such as Fig. 1: (a) for five different dichotomies, (b) for two trichotomies and two dichotomies. Label the cells and show that the figure exhausts the possibilities by listing them.
3. Draw up a classification scheme for dividing the set of all prices into subsets and illustrate by a diagram such as Fig. 1.
4. What are the dimensions, in terms of the length-mass-time system, of the following physical quantities: (a) atmospheric pressure, (b) rainfall, (c) stream flow, (d) density.
5. What are the dimensions, in terms of the commodity-money-time-length system, of the following economic quantities: (a) gold reserves, (b) depreciation of capital, (c) yield of crops per acre, (d) density of population, (e) elasticity of demand, (f) a specific tax, (g) an ad valorem tax. (Note: in determining the dimensions of any quantity, always start by investigating the nature of the unit in which it is measured.)
6. List the economic quantities which have the following dimensions:

$$\text{(a)}\ \frac{\text{Commodity I}}{\text{Commodity II}} \qquad \text{(b)}\ \frac{\text{Money}}{\text{Length}} \qquad \text{(c)}\ \frac{\text{Money}}{\text{Time}}$$

7. Suppose that it takes 10 bushels of wheat to produce 8 bags of flour, and that 1 bag of flour produces 30 loaves of bread. Suppose that in a given time period gross production is 120 bushels of wheat, 88 bags of flour, and 2100 loaves of bread. Net consumption (that is, consumption not resulting in further production) is 5 bushels of wheat, no flour, and 2100 loaves of bread. What is the net production, the gross consumption, and the total accumulation of each of the three commodities? Show that the accumulation can be expressed either as the difference between gross production and gross consumption, or as the difference between net production and net consumption in each case.

CHAPTER 2

THE MEASUREMENT OF ECONOMIC AGGREGATES

The problem of the measurement of economic aggregates is not confined to macroeconomics; we run into it, for instance, in accounting theory, in discussing the measurement of the aggregates of the balance sheet (see Volume I, Chapter 15). Even in private accounting this problem becomes acute when the valuation coefficients are not stable over time or space, for unique solutions to the problem can only be given when these coefficients are stable. Exactly the same problem arises in national accounting, whether this is in the computation of aggregate balance sheets and position statements or in the computation of national income and product.

Measurement of Aggregates of Goods

The first problem is how to measure the physical quantity of some heterogeneous aggregate—which may be an aggregate of stocks, as in the balance sheet, or an aggregate of flows, as in the income statement. We cannot obtain such a measure by simply adding the various quantities in the inventory, for the numbers representing these quantities are quite arbitrary, depending on the size of the unit of measurement. Thus an identical quantity of wheat may be represented by the number 1 if measured in short tons, 2,000 if measured in pounds, 32,000 if measured in ounces, and so on. The incommensurability of a stock of various goods may be seen clearly if we ask ourselves this question: "Add a grand piano to the total stock and take away a cow; has the total stock increased or decreased?" There is clearly no way of answering this question unless we have some measure of assessing the significance or value of the various items in the inventory. If a grand piano is equivalent in some way to three cows, then adding a grand piano and subtracting a cow is clearly

equivalent to adding two cows, or ⅔ of a piano. If, however, a grand piano is only equivalent to half a cow, then adding a grand piano and subtracting a cow is equivalent to subtracting half a cow or one piano. It is evident that a list or inventory of things can only be summed if each quantity can be converted into some common equivalent by multiplying it by some valuation coefficient, or ratio of equivalence. This "ratio of equivalence" is called in statistics the "weight" of the quantity in question. This commonest weight is, of course, the money price of the commodity. If the physical quantity of each commodity is multiplied by its price, a dollar value is obtained; these dollar values can be added together, as in Table 1, to get the total value of the inventory. It is

TABLE 1. AGGREGATES OF GOODS

Commodity	Quantity	Price	Value
Wheat	50 bushels	$ 1.80 per bushel	$ 90
Eggs	8 dozen	.50 per dozen	4
Shirts	30 shirts	4 00 per shirt	120
Drills	3 drills	100.00 per drill	300
			$514

clear in this example that adding bushels, dozens, shirts, and drills makes no kind of sense. Multiplying each quantity, however, by its price gives us a list of values, each of which is expressed in the same unit—the dollar—and which can be added to get a sum of dollar values, or *dollars' worth.*

INDEX NUMBERS

Ambiguity in Indices

This is essentially the procedure by which balance sheets and income accounts are constructed, as we have seen. It must be emphasized, however, that the dollar values of an inventory of stock, or of those changes in stock which constitute income in a physical sense, are only significant as measures of physical volume if the prices at which the items are valued are constant and in some sense significant. If relative prices are not constant, there is an unavoidable ambiguity in the very concept of a change in the quantity of goods, for the result will depend on whether we take prices prevailing at the beginning or at the end of the period under consideration. The point can perhaps best be illustrated by an example, as in Table 2. Here we take two commodities only for the sake of simplicity—wheat and shirts. The quantity at the beginning of the change (the base date) is indicated by the column

q_0, the quantity at the end of the change (the end date), by the column q_1; the corresponding prices are p_0 and p_1.

On the base date the total value of wheat and shirts together (V_0) is \$210; on the end date the total value (V_1) is \$310. This change, though it is significant for many purposes, does not measure in any sense the change in the physical quantity of wheat and shirts; for the weights (i.e., the prices) have changed. It is possible, for instance, for the total

TABLE 2. PRICE AND QUANTITY INDICES

	q_0	p_0	q_0p_0		q_0	p_1	q_0p_1
Wheat	50	\$1.80	\$ 90	Wheat	50	\$3.00	\$150
Shirts	30	4.00	120	Shirts	30	2.00	60
Total (V_0)			\$210	Total (W_0)			\$210
	q_1	p_0	q_1p_0		q_1	p_1	q_1p_1
Wheat	100	\$1.80	\$180	Wheat	100	\$3.00	\$300
Shirts	5	4.00	20	Shirts	5	2.00	10
Total (W_1)			\$200	Total (V_1)			\$310

value of a stock of goods to increase merely because of increase in prices without any change whatever in the composition of the stock. To obtain any measure of the change in physical quantity we must compute either what would be the value of the stock on the end date at base-date prices, W_1, and compare this figure with the value of the stock at the base date at base-date prices, V_0; or we must calculate what would have been the value of the stock on the base date at end-date prices, W_0, and compare this with the actual value at the end date at end-date prices, V_1. In the example the figures have been chosen for convenience to make the base-date values the same both at base-date and at end-date prices $(V_0 = W_0 = \$210)$. In this case, however, we find that valued at base-date prices the total stock has fallen from \$210 to \$200, whereas valued at end-date prices the same stock has risen from \$210 to \$310!

In this particular case we emerge with the paradoxical result that measured in one way the total quantity of goods has fallen, and measured in another way the total quantity has risen—the detailed changes in both cases being identical! There is no escape from this dilemma; there can be no single definition of a "total quantity of goods," because the total quantity can only be defined on the assumption of some system of prices or ratios of equivalence, and the system of prices which may be relevant for one purpose may not be relevant for another.

It must not be thought, however, that because of this difficulty the concept of a total physical stock of goods is without meaning. In any particular discussion the concept can be defined as accurately as we wish by defining the relevant system of prices. Moreover, in fact, relative quantities or prices do not usually shift radically, especially over fairly short periods, so that fairly similar results are obtained no matter whether base-date, end-date, or any other reasonable prices are used as weights. If the relative quantities are unchanged (that is, if all quantities change in the same proportion), the quantity index is independent of the price weights used. Thus if all quantities double, the index will double, no matter what prices are used.[1]

Measurement of Price Levels

The problems which are involved in measuring price levels are essentially similar to those involved in measuring quantity aggregates. The problem is to get some single figure which is descriptive of change in a given aggregate or list of prices. If all prices change in the same ratio there is no problem—any measure of the average of all the prices, or the price level, will change in the same ratio. If prices change in different ratios, however, the problem of how to weight the various changes arises, just as in the case of quantity aggregates.

Price Levels and Quantity Levels

The close relationship between the measurement of price levels and of quantity aggregates can be seen if we define the price level in the same way as we define a price—as a ratio of a quantity of money to the quantity of goods which is exchanged for the money. Thus the price of cheese is the ratio of the quantity of money paid for cheese to the quantity of cheese bought. If $100 is paid for 200 pounds of cheese, the price is 50 cents per pound.

[1] Suppose that for any commodity, $q_1 = kq_0$, k being a constant.

Then $V_0 = \Sigma p_0 q_0$ $\qquad W_0 = \Sigma p_1 q_0$

$W_1 = \Sigma p_0 q_1 = k\Sigma p_0 q_0$ $\qquad V_1 = \Sigma p_1 q_1 = k\Sigma p_1 q_0$

Then the quantity index at base-date prices, $Q_b = \dfrac{W_1}{V_0} = k$, and the quantity index at end-date prices, $Q_e = \dfrac{V_1}{W_0} = k$. The quantity index is always k no matter what the prices used.

Similarly, if relative prices are constant, for any commodity $p_1 = Kp_0$.

Then $W_0 = K\Sigma p_0 q_0 = KV_0$ and $V_1 = K\Sigma p_0 q_1 = KW_1$

and $$Q_1 = \frac{V_1}{W_0} = \frac{KW_1}{KV_0} = \frac{W_1}{V_0} = Q_0$$

The quantity index is independent of the actual *level* of prices but depends only on their relative structure.

Similarly, a price level of a given list of commodities is the ratio of the total money value of the commodities to their total quantity; that is,

$$\text{Price level} = \frac{\text{total money value of commodities}}{\text{total physical quantity of commodities}}$$

All the difficulties and ambiguities, therefore, which we noted in considering the concept of the total quantity of a list of heterogeneous commodities apply equally to the concept of a price level. Just as the way to measure a change in a quantity level is to calculate the value of the list of commodities at different dates on the assumption of constant prices, so the way to measure a change in the price level is to calculate the change in the value of a constant list or "basketful" of commodities. Thus, in Table 2 the change in price level between the two dates may be measured either by comparing the value of the base-date quantities first at base-date and then at end-date prices—i.e., by comparing W_0 with V_0; or by comparing the value of end-date quantities, first at base-date and then at end-date prices—i.e., by comparing V_1 with W_1. Again the ambiguity shows up. The value of the base-date quantities, 50 bushels of wheat plus 30 shirts, has not changed at all between the two dates; the value of the end-date quantities (100 bushels of wheat plus 5 shirts) has increased from \$210 to \$310. On the first criterion there has been no change in the price level. If the index of the base date is 100, the index of the end date will also be 100. In the second case, however, there has been apparently a substantial increase. If the index of the base date is 100, that of the end date is $\frac{310}{210}(100)$ or 147.6.

The Index as a "Range"

There is no particular sense in asking which of these two apparently contradictory results is "right." Each of them is meaningful, in the sense that it has a simple physical interpretation. In some sense these values constitute the extreme limits of the meaningful range of values. There is much to be said for quoting the extreme range in using these indices. Thus we might say in the above case that the price level has moved from 100 to between 100 and 147.6. This statement gives us as much information as the nature of the problem entitles us to have. It is important information: we know that the price level has not doubled, and has not fallen. It would no doubt be agreeable to know that it has risen to exactly, let us say, 125, but in the nature of the problem such a figure would be meaningless. It should be observed that this problem of the "range" is present in all measurement.

When we say that something is 10 inches long, we mean strictly that measured with instruments capable of distinguishing hundredths of an inch it is between 9.99 and 10.01 inches. If the ranges are broader in the case of economic aggregates, the apparent inaccuracy is merely a matter of degree, and is inherent in the nature of the problem.[2]

How Significant Are Indices?

It is evident that the significance of a price index depends on the retention of significance of the list of commodities comprised in it. For this reason comparisons of price levels are more and more difficult the longer the period of time over which the comparisons are made. Even over fifty years the composition both of the stock of goods and of the flows of production and consumption change enormously. How can we compare, for instance, the automobiles and washing machines of today with the "horseless carriages" and washtubs of 1900? Still more difficult is it to make any meaningful comparison between the goods of today and the chariots and fibulas of ancient Rome. It is for this reason that the measurement of price levels over long periods is confined for the most part to the staple wholesale commodities—grains, metals, and so on—which remain much the same from century to century. Even cost of living indices (that is, price levels of consumers' goods) become almost meaningless in a few decades.

Consistency of Indices

For all their unsatisfactory nature, the price-level, quantity-level concepts cannot be dispensed with, and the temptation to deal only with measurable dollar values must be resisted. The dollar value of any commodity is a product of a price and a quantity, and similarly the dollar value of a group of commodities is in some sense a product of a price level multiplied by a quantity level. The exact division of any given dollar value into its two factors may be in some degree

[2] This basic conceptual difficulty vitiates the various attempts which have been made (notably that of Irving Fisher in *The Making of Index Numbers,* New York, 1923) to construct an "ideal" index number which would give an unequivocal figure for a price level. There are indices (e.g., the weighted geometrical means) which are superior to the weighted arithmetic indices used above as judged by certain statistical tests, such as reversibility. (Reversibility means giving the same result whether calculated forward or backward in time.) There are no other indices which have such a simple and obvious economic meaning, however, as the weighted arithmetical indices which measure the change in the value of a given basketful of goods. There are no other indices, either, which are so easy to compute. For these reasons the weighted arithmetical index is predominant. It is usually weighted by base dates, as the evaluation of the weights is a difficult and expensive task and is not repeated annually. Cost of living indices are almost universally arithmetic averages weighted by quantities derived from budget studies at the base date.

arbitrary, but this arbitrariness is implicit in the very nature of the problem. As long as our concepts of price level and quantity level are consistent, the element of arbitrariness need not cause any confusion of thought. The test of consistency is that the product of the price-level index and the quantity-level index should be equal to the index of the change in dollar value.

Following the notation of Table 2, we have:

$V_0 = \Sigma p_0 q_0 =$ the value of base-date quantities at base-date prices
$W_0 = \Sigma p_1 q_0 =$ the value of base-date quantities at end-date prices
$W_1 = \Sigma p_0 q_1 =$ the value of end-date quantities at base-date prices
$V_1 = \Sigma p_1 q_1 =$ the value of end-date quantities at end-date prices

V_0 and V_1 are actual dollar values; W_0 and W_1 are fictitious. From these we derive four possible concepts of price and quantity indices, or relatives, for the end-date with the base-date index = 1.

$$P_b = \frac{W_0}{V_0} = \text{price level with base-date quantity weights}$$
$$P_e = \frac{V_1}{W_1} = \text{price level with end-date quantity weights}$$
$$Q_b = \frac{W_1}{V_0} = \text{quantity level with base-date price weights}$$
$$Q_e = \frac{V_1}{W_0} = \text{quantity level with end-date price weights}$$

It is evident immediately that P_e and Q_b are consistent measures, for $P_e Q_b = \frac{V_1}{V_0}$, the index of the change of real values. Similarly, P_b and Q_e are consistent measures. P_b and Q_b, or P_e and Q_e would not be consistent. The rule, therefore, emerges that if end-date price weights are used in calculating the quantity index, base-date quantity weights must be used in calculating the price index, and vice versa.

Graphic Analysis

The principles underlying the use of index numbers can perhaps be made clearer with the aid of the following graphical analysis. For simplicity we assume two commodities, one of which is money. In Fig. 3, then, we measure the quantity of money along the vertical axis, and the quantity of some commodity, say wheat, along the horizontal axis. Any point in this field then represents a combination of quantities of money and wheat. These two quantities can be regarded as stocks or as flows—it makes no difference to the argument. Suppose then that we have two money-wheat combinations, A_0 with OH_0 wheat

and H_0A_0 money, and A_1 with OH_1 wheat and H_1A_1 money. The problem is which of these two combinations is the larger, and by how much. The question cannot be answered, as we have seen, unless we have a system of valuation ratios—i.e., unless we can express any quantity of the one asset as an equivalent quantity of the other. In this example this means we must know the price of wheat, as the price

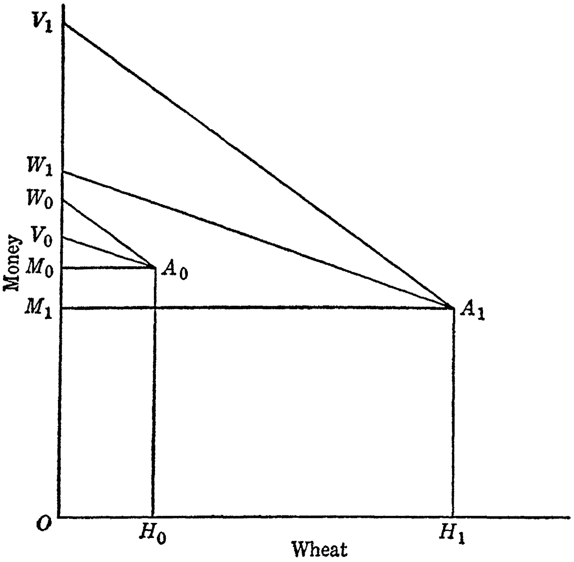


Fig. 3. Price and Quantity Indices

of money is 1. The price of wheat can be expressed by the slope of a line drawn from any point such as A_0 or A_1. Thus suppose we draw a line A_0V_0 to the vertical axis, and drop a perpendicular to the same axis A_0M_0. Then if M_0V_0 is the money value of the amount of wheat M_0A_0 ($= OH_0$), the price, or more strictly the valuation coefficient, of wheat is $\frac{M_0V_0}{M_0A_0}$, which is the slope of the line A_0V_0, or the tangent of the angle $V_0A_0M_0$. Thus if M_0V_0 represented \$100 and M_0A_0 represented 50 bushels, the price of wheat would be $\frac{100}{50}$, or \$2 per bushel. The line A_0W_0 represents a higher price of wheat, giving a higher money value (M_0W_0) for the same quantity, M_0A_0. Suppose now that the slope of the line A_0V_0 shows the base-date price of wheat at the date when the quantity combination is at A_0, and that the slope of A_0W_0 shows the end-date price of wheat at the date when the quantity combination is at A_1. OV_0 is the total money value of the money and wheat together

at A_0, at base-date prices (OM_0 is the value of the money, M_0V_0 the value of the wheat). Similarly OW_0 is the total money value of the A_0 combination at end-date prices. If we draw A_1W_1 parallel to A_0V_0, the slope of A_1W_1, being the same as the slope of A_0V_0, measures the base-date price of wheat, and OW_1 is the total money value of the A_1 combination at base-date prices. Similarly drawing A_1V_1 parallel to A_0W_0 gives us the end date or actual value of the A_1 combination, OV_1.

Value Change as a Combination of Price and Quantity Changes

The movement from OV_0 to OV_1 represents the increase in the total money value, at current prices, represented by moving from A_0 to A_1. The movement is partly an increase in quantity, partly an increase in price. The breakdown between the quantity and the price movement can be made either at W_0 or at W_1. If it is made at W_0, then V_0W_0 measures the price change at base-date (A_0) quantities. This is the change in the value of the basketful represented by A_0. W_0V_1 measures the quantity change at end-date prices. These are consistent indices, as V_0W_0 and W_0V_1 together give the total change in value, V_0V_1. Similarly we may make the break at W_1, in which case V_0W_1 shows the quantity change at

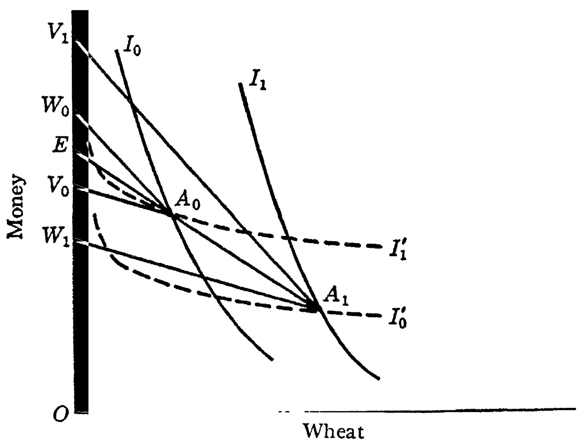


Fig. 4. Price and Quantity Indices

base-date prices, and W_1V_1 shows the price change at end-date quantities. It is not consistent, however, to take, say, V_0W_0 as the measure of price change and V_0W_1 as the measure of quantity change, as these do not exhaust the total movement in current value.

In Fig. 4, which is essentially similar to Fig. 3, we see how a given change can be interpreted either as a rise or as a fall in physical quantity, depending on the price weight, or valuation coefficients used. Thus at a

high price of wheat, represented by the slopes of A_0W_0 or A_1V_1, the position A_1, with more wheat and less money, represents a larger total (OV_1) than A_0 (OW_0). At a low price of wheat, represented by the slopes of A_0V_0 or A_1W_1, the position A_0, with little wheat and a lot of money, is a larger total (OV_0) than A_1 (OW_1). There is also some set of valuation coefficients or weights which will make any two quantity sets equal in value: thus in Fig. 4 if the price of wheat is equal to the slope of the line A_0A_1, the value of both combinations is OE.

What we are really asking ourselves is not really which of two quantity sets is the *bigger* but which is the *better*. This problem can be solved conceptually by means of indifference curves (see Volume I, Chapter 27). Thus in Fig. 4, if the indifference curve through A_1, I_1 is superior to that through A_0, I_0, then A_1 is the superior quantity set. With indifference curves such as the dotted curves I'_0 and I'_1, however, A_0 would be the superior position. What we are really doing when we assume a given set of price weights or valuation coefficients is to suppose that the lines of equal value at these weights, such as the whole set of lines parallel to, say, A_1V_1 or A_0W_0, are in fact indifference curves—that is, that we are indifferent to any combination having the same total value and always prefer a higher total value to a lower. This is a highly restrictive assumption; there is no real reason to suppose, for instance, that the preference function will be linear in this sense. However, we must be content with this in practice, simply because even if we could trace indifference curves by questionnaire, which is rather unlikely, we could not number them with more than an ordinal numbering, and hence could not derive a cardinal index or weight from them.

NATIONAL INCOME ACCOUNTING

One of the most important developments in economics since the 1920s has been the rise of national income accounting by governments. Before 1929 there was no good information about the gross aggregative structure of any economy. The collection and publication of information regarding national income and product, its sources and its distribution, have resulted in an impressive cumulation of knowledge about the most important economies of the world. The facts so revealed are not only vital to the intelligent pursuit of economic policy but have also forced economists to clarify concepts and to fit their theories into a form which is consistent with the collection of aggregate information. The collection of this information has also led to the debunking of many economic myths, which could be held without possibility of contradiction only so long as no adequate information was available by which they might be tested. Thus the Marxist theory of increasing misery of the workers looks

foolish in the presence of national income statistics which show the labor force receiving a rather constant share of a rapidly increasing "pie" of national income. Henry George's idea of the "single tax" fares ill when we compare what is absorbed by government with the small fraction of national income which goes to rents and royalties.

National income accounts are quite complex, and before going on to the actual accounts, for instance, of the United States, the following simplified version, involving two or three concepts which are close to but not quite identical with the concepts of the standard national accounts, may be helpful. As the most easily available source of the United States national income accounts is in the tables at the back of the *Economic Report of the President,* appendix B, we shall use a notation which gives quick reference to the figures involved. I have used the January 1965 edition, the latest available at the time of writing. Columns in each table are lettered alphabetically, then a notation, such as D_7, refers to the fourth column of table 7. Concepts not covered in these tables are indicated with a prime ($'$). A simplified system of national income accounts is shown in Fig. 5.

The key to the understanding of these accounts is the realization that they start with product, and that income is always derived from the *value* of a product of goods and services. The gross national product, A_1, then, often abbreviated to GNP, is the dollar (money) value of goods and services produced by the people of a nation in a given period of time, usually a year, double counting being eliminated with the exception of the depreciation of fixed capital.

The GNP can be broken down into segments in many ways. We can break it down first into the various ways of disposing of the product. Four major categories of disposal are usually recognized. The first is usually called personal consumption expenditures, B_1. It is not the expenditures which are significant here, but the value of the goods purchased and taken off the market, which is the same as the amount of money spent for them. Thus we include household receipts in kind, and presumably subtract household sales of secondhand goods to businesses. The second is called gross private domestic investment, C_1. This consists of the increase of the value of goods of all kinds in the hands of businesses, net private domestic investment, C_{1n}, plus what is called the capital consumption allowance, B_{12}. Perhaps the simplest way to justify this concept is to think of gross private domestic investment as the total value of "new goods" added to the stock of goods held by businesses, which should be included in the gross national product, but which is offset to some extent by the depreciation in value of "old goods" already in existence at the beginning of the year.

The third segment of the GNP is called net exports of goods and

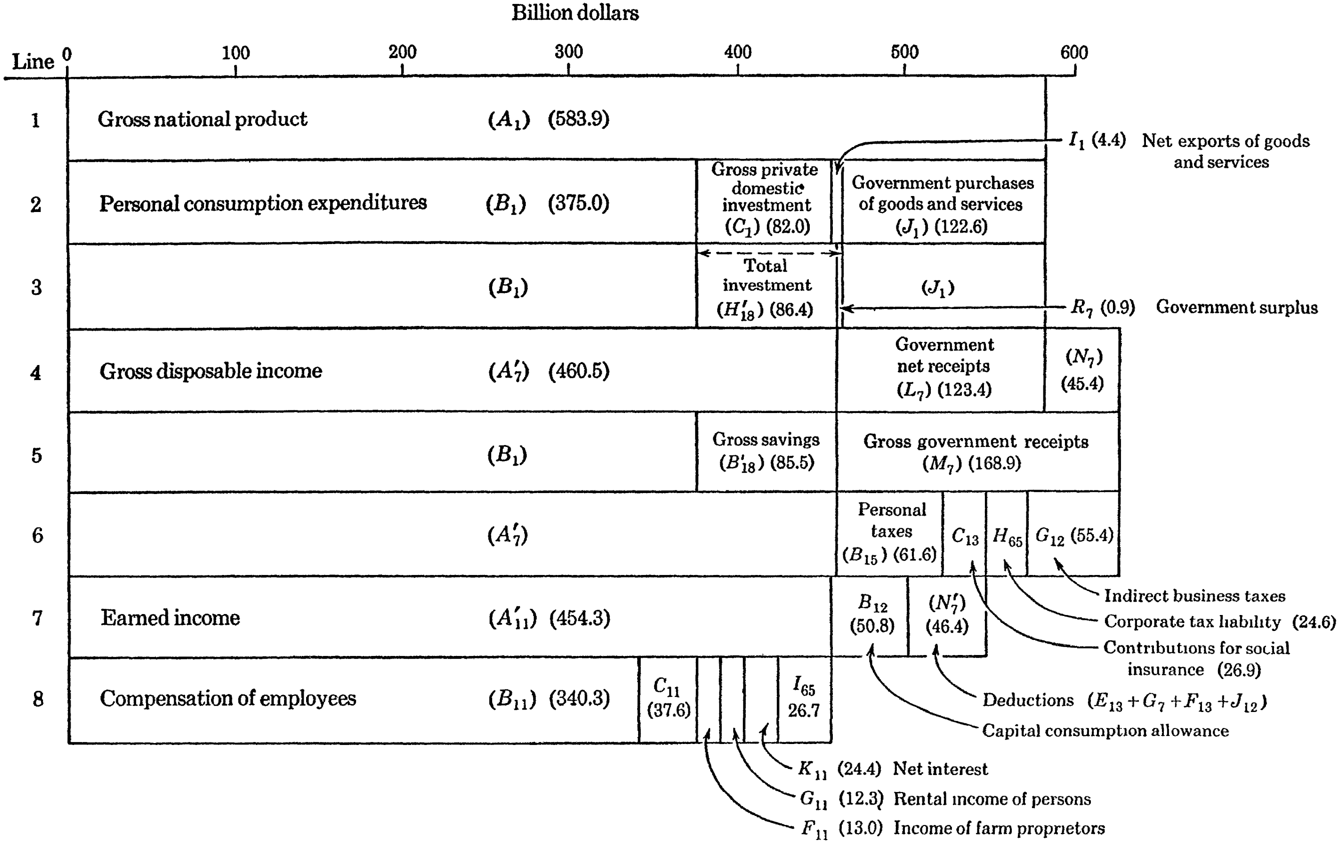


Fig. 5. U.S. National Income Accounts, 1964: A Simple Model

services, I_1, and is equal to exports minus imports. Exports diminish the stock of goods at home, imports increase the stock; so that net exports represents a net diminution in the stock. The fourth segment is total government purchases of goods and services, J_1, which includes the purchase of labor and personal services from government employees and net purchases of goods from the rest of the economy.

The first equation, or rather identity, of the national income accounts can then be written

$$A_1 = B_1 + C_1 + I_1 + J_1 \tag{1}$$

It is shown in lines 1 and 2 in Fig. 5. What this means can be visualized most clearly if we suppose that the GNP all originates in the hands of businesses; then some of it is disposed of to households, B_1, some to government, J_1, and some as net exports abroad, I_1; what remains is still in the hands of businesses and appears as an addition to the value of goods held by them, C_1, not counting the capital consumption of old goods.

In line 3 of Fig. 5 we give a slightly different breakdown; we combine I_1 and C_1 in a concept we can call total investment, H'_{18},—not found in this exact form in the United States accounts—and we can break down government purchases into net government receipts, L_7, and the government deficit, $-R_7$, called government surplus or deficit $(-)$ on income and product account in the United States national accounts. We shall find it more convenient to treat the deficit as positive, rather than as negative, as is done in the United States national accounts. We have, therefore, two more identities:

$$H'_{18} = C_1 + I_1 \tag{2}$$

$$J_1 = L_7 + (-R_7) \tag{3}$$

In lines 4 and 5 we see how the net government receipts is equal to *gross government receipts*, M_7—*taxes and nontax receipts or accruals (government)*, in the United States accounts—minus government transfers, N_7, or

$$L_7 = M_7 - N_7 \tag{4}$$

In line 4 we define a new concept, which I call gross disposable income, A'_7, defined simply as GNP less net payments from the private sector to government: that is,

$$A'_7 = A_1 - L_7 \tag{5}$$

We now define gross saving, B'_{18}, as the difference between gross disposable income and personal consumption expenditures. This is not a wholly satisfactory definition of saving, which is more refined in the United States national accounts; however, it makes some sense and has the virtue

of illustrating simply a very important identity of the accounts known as the "sources and uses of savings" identity, of which the famous "savings = investment" identity of Keynes is a special case. In this case we see immediately from the figure that gross savings is equal to total investment plus the government deficit. Algebraically we have:

$$\begin{aligned} B'_{18} &= A'_7 - B_1 && (6) \\ &= A_1 - L_7 - B_1 \\ &= C_1 + I_1 + J_1 - L_7 \\ &= H'_{18} + (-R_7) && (7) \end{aligned}$$

To appreciate the full significance of this identity we must look at it from the point of view of the capital account. Income in its various forms and definitions always implies a gross increase in net worth. Consumption implies a gross decrease in net worth. Saving, therefore, which is the excess of income over consumption, is equivalent to a net increase in aggregate net worth. Net worth can only increase, however, if there is a corresponding increase in net assets; net assets can increase either because of an increase in the value of real capital (total investment) or because of a budget deficit which increases the total value of the obligations of government held by private persons, either in the form of government securities (national debt) or in the form of new money. This is why a budget deficit *creates* saving if it is not offset by a decline in total investment.

The gross national product and its various derivatives can be distributed not only according to the various methods of disposing of it, but it can also be allocated to the various factors of production and other recipients which have contributed to the product. This gives us an alternative method of computing the GNP, which provides something of a check on the calculations. The difference in the sums produced by the two methods is called the statistical discrepancy, T_7, and is rarely more than 0.5 percent of the total in the United States accounts. Here again we can illustrate the process most simply by means of a concept which differs slightly from those used in the United States accounts. We will define earned income, A'_{11}, as the sum of income (before paying personal taxes), B_{15}, and personal contributions for social insurance, C_{13}, received by the various factors and groups contributing to production. It is the sum of corporate profits after taxes,[3] I'_{63}, compensation of employees, B_{11}, business and professional income (unincorporated)[3], C_{11}, income of farm proprietors, F_{11}, rental income of persons, G_{11}, and net interest, K_{11}. That is:

$$A'_{11} = I'_{65} + B_{11} + C_{11} + F_{11} + G_{11} + K_1 \qquad (8)$$

[3] These include an allowance for revaluation of inventory.

This is shown in lines 7 and 8 of Fig. 5. Not all the income which is generated by producing the GNP, of course, is allocated to earned income. Some goes to the capital consumption allowance, B_{12}. Some goes to pay indirect business taxes, G_{12}, and corporate taxes, H_{65}, for these taxes represent an increase in the prices of the final product beyond the cost of the product at factor prices plus profits. Finally a small adjustment must be made for business transfers, J_{12}, which represent mostly the bad debts of businesses, which are income not counted in factor payments to persons who incur these debts; we can think of this as a kind of informal addition to business taxes, the assumption being that this is taken care of in the price of the product. An allowance must also be made for another small item: subsidies less current surplus of government enterprises, F_{12}, so that we have finally

$$A_1 = A'_{11} + B_{12} + G_{12} + H_{65} + J_{12} - F_{12} + T_7 \tag{9}$$

The significance of this identity may perhaps be seen more clearly in terms of Fig. 5 if we add the item N_7, government transfers, interest, and subsidies, to both sides of the equation. The item N_7 itself is made up of four components: government transfer payments to persons, E_{13}, foreign net transfers by government, G_7, net interest paid by government, F_{13}, and the net-government-subsidies item mentioned above, F_{12}. That is:

$$N_7 = E_{13} + G_7 + F_{13} + F_{12} \tag{10}$$

We have therefore

$$A_1 + N_7 = A'_{11} + B_{12} + (E_{13} + G_7 + F_{13} + J_{12}) + H_{65} + G_{12} \tag{11}$$

We can write the item in brackets:

$$N'_7 = E_{13} + G_7 + F_{13} + J_{12} \tag{12}$$

and regard it as the transfer income of persons not counted in earned income but generated by the GNP plus government transfers.

Equation (11) is then shown by comparing lines 4 and 7 of Fig. 5.

Criticism of National Income Accounts

A system of national accounts of the above type is open to a number of criticisms from a theoretical point of view, though it is not always easy in practice to compute accounts according to any theoretical ideal; the accounts are always a compromise between what we can find out and what we want to know. Perhaps the most serious criticism of the accounts is their failure to integrate income and capital accounting. What the national

income accounts call personal consumption expenditures is not consumption, but household purchases plus household production: it is equal to real household consumption plus household investment, or accumulation of goods by households. Household capital in these days is an important item, and so is household investment; a washer and dryer in the kitchen is just as much capital as a machine in a factory and may have an even larger effect on overall productivity! Similarly it would be very desirable to divide government purchases into consumption and government investment; the situation of the economy when government is building roads and dams is very different from the situation when government is engaged only in current consumption. The national income accounts thus seriously underestimate both real investment and saving.

A defect in national accounts which could perhaps be rectified is the omission of family labor, especially of housewives, mainly on the grounds of difficulty of estimation. This leads to the paradox that if a respectable man marries his housekeeper the national product falls, as services which previously were in the market sector now fall within the confines of the household; one would think in reality that this happy event should make the national product rise!

Many difficult conceptual problems relate to the role of government. In computing the GNP it is assumed that government purchase of goods and services represent a "final" product; this treats the government almost as if it were a person or household. At the other extreme we might argue that government is simply a kind of overhead which is a cost of producing the final "private" product, and that something like net disposable income (gross disposable income less the capital consumption allowance) would be a better measure of final product. The truth probably lies between these two extreme positions, but it is virtually impossible to divide the government purchases into what should go to final product and what to overhead. Thus even national security expenditures can be regarded either as mere overhead, a cost which is incurred to maintain the private economy intact against possible disrupters from abroad, or it can be regarded as a kind of public consumer good which makes the bosom of the citizen swell with pride as he contemplates the might of his nation.

Problems which seem to be inherently more soluble but in which compromise has to be made because of the difficulties of collecting information lie in the field of allocating taxes and government transfers, which may be thought of as negative taxes. The distribution of income by functional shares would be more meaningful if we could allocate to each share a proportion of the personal taxes paid; this, however, seems to be difficult in practice.

National Income Accounts in Real Terms

For many purposes we wish to know not only the values of the various items in the national income accounts in current dollars, but their value in real terms, that is in dollars of constant purchasing power. To do this an *implicit price deflator* must be calculated for each item of the accounts, based on the price index which is most relevant to each sector. Thus the implicit price deflator for personal consumption expenditures for 1964, with 1954 = 100 is 120.7; this means that the price index of goods entering into personal consumption was 1.207 times higher in 1964 than in 1954. Then to reduce the 1964 current dollar figure for personal consumption expenditures into 1954 dollars we divide it by 1.207. Care must be taken that the various implicit price deflators are compatible, so that the basic identities are still met. Thus to turn Fig. 5 into the corresponding figure in 1954 dollars we shrink each item by an appropriate amount; if we are given the implicit price deflators for, say, the four components of GNP in line 2, this automatically yields us the implicit price deflator for the GNP itself in line 1, as even in constant dollars the GNP must equal its four major components.

QUESTIONS AND EXERCISES

1. What different meanings can be attached to the word "measurement"? Are the problems of measurement essentially different in the social sciences from what they are in the physical sciences? If not, why does measurement seem to be more difficult in the social sciences?
2. Three students, X, Y, and Z, take examinations in three subjects with the following results:

	X	Y	Z
Math	D	A	B
English	A	F	B
Biology	B	A	B

Compare the total marks of the three students under each of the four following systems of grade weights and comment on the results:

System	*Grade*				
	A	*B*	*C*	*D*	*F*
(1)	100	90	80	70	60
(2)	100	99	98	97	96
(3)	100	30	20	10	0
(4)	100	65	55	45	0

What analogies might you draw with the problem of price and quantity indices?

3. Figure 4 shows how different price weights can measure a given change in quantities either as an increase or decrease in the aggregate. Construct a similar figure to show how different quantity weights can measure a given change in

prices either as an increase or decrease in the price level. (Note: By buying wheat futures, we can have a negative quantity of wheat!)

4. The following table shows the prices and sales of three commodities at two different localities, X and Y.

	Locality X		*Locality Y*	
Commodities	*Price*	*Quantity*	*Price*	*Quantity*
A	100	50	80	5
B	5	1000	10	10
C	20	30	15	2

Calculate (a) an index of the physical volume of sales in the two localities for each set of prices; (b) an index of the price level of the sales in the two localities for each set of quantities; (c) an index of the total value of sales in each locality.

In calculating all the indices make the value in locality Y equal to 100. Which of these sets of indices are consistent and which are not? Comment on the results.

5. Redraw Fig. 5 with the net national product, E_{12}, that is, the gross national product, A_1, less the capital consumption allowance, B_{12}, on the top line, and all the other lines and concepts adjusted accordingly.
6. Using the notation of this chapter, prove the following identity;

$$A'_7 + B_{15} + C_{13} = A'_{11} + T_7 + B_{12} + N'_7$$

Illustrate it from Fig. 5 and comment on its economic significance, if any.

APPENDIX

THE NATIONAL INCOME ACCOUNTS OF THE UNITED STATES

There are a good many ways in which national income accounts can be computed, and in practice these accounts are more complicated than the simple model of this chapter. Figure 6 illustrates the national income ac-

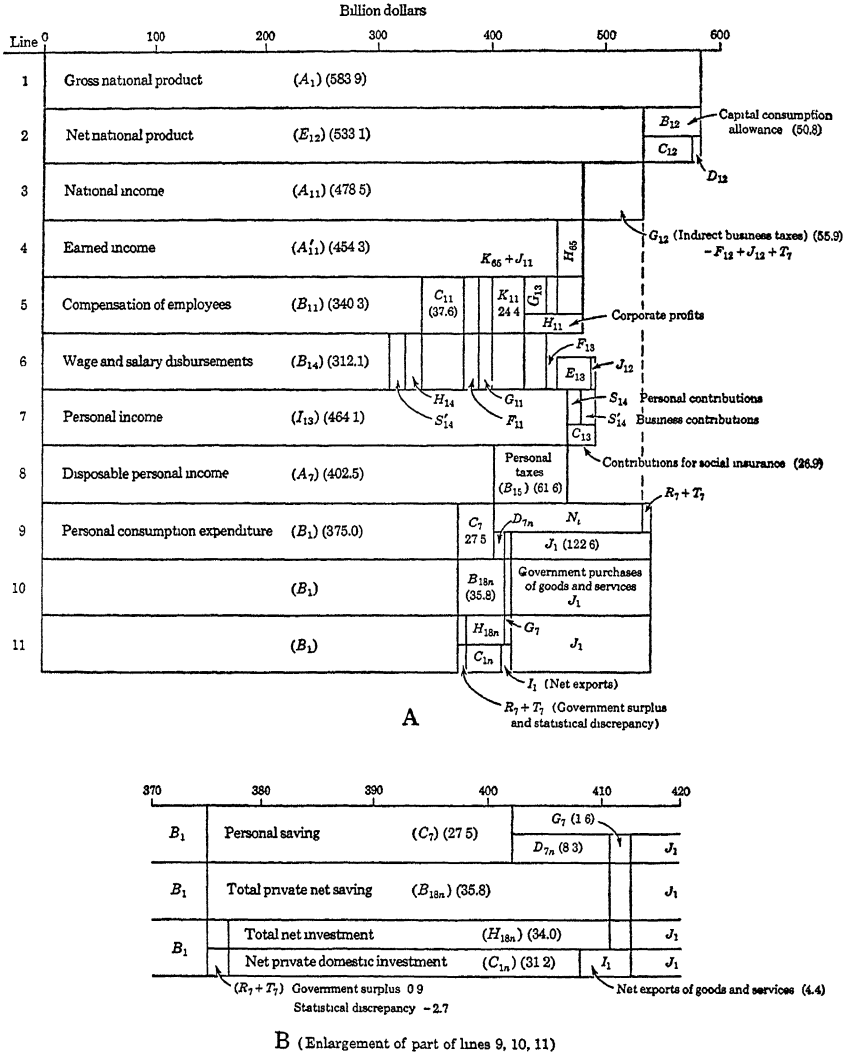


Fig. 6. U.S. National Income Accounts, 1964

counts of the United States for 1964. Again we will use a notation which refers to the tables in appendix B, the *Economic Report of the President*, January, 1965. Then in line 1 we have the GNP, as before. In line 2 this is broken down into the net national product, E_{12}, and the capital consumption allowance, B_{12}. The capital consumption allowance in turn is broken down into two parts, depreciation charges, C_{12}, and other capital consumption allowances, D_{12}, which mainly consist of losses due to fire, etc. Line 3 then shows how the national income A_{11}, is derived from the net national product by adding a small item entitled subsidies less current surplus of government enterprises, F_{12}, and subtracting indirect business taxes, G_{12}, business transfer payments, J_{12}, and the statistical discrepancy, T_7. We have, that is,

$$E_{12} = A_1 - B_{12} \tag{1}$$

$$A_{11} = E_{12} + F_{12} - G_{12} - J_{12} - T_7 \tag{2}$$

Line 4 shows the relation of the national income concept of the United States accounts to the earned income, A'_{11}, of page 35; earned income equals nation income, less corporate tax liability, H_{65}, that is,

$$A'_{11} = A_{11} - H_{65} \tag{3}$$

Lines 3 and 5 now show how the national income is built up out of the income of factors. It consists of an item, H_{11}, called corporate profits and inventory valuation adjustment, plus compensation of employees, B_{11}, plus an item called business and professional income and inventory valuation adjustment, C_{11}, plus income of farm proprietors, F_{11}, plus rental income of persons, G_{11}, plus net interest, K_{11}. That is,

$$A_{11} = H_{11} + B_{11} + C_{11} + F_{11} + G_{11} + K_{11} \tag{4}$$

Of these components, H_{11} can be broken down into corporate tax liabilities, H_{65}, undistributed corporate profits, K_{65}, the corporate inventory valuation adjustment, J_{11}, and dividends G_{13}, that is,

$$H_{11} = H_{65} + K_{65} + J_{11} + G_{13} \tag{5}$$

We also have in the accounts an item, I_{11}, corporate profits before taxes $(H_{11} - J_{11})$; and I_{65}, corporate profits after taxes $(G_{13} + K_{65})$. The item I'_{65} of equation (8) (p. 35) is $G_{13} + K_{65} + J_{11}$, so that $H_{11} = H_{65} + I'_{65}$, as we see in lines 4 and 5. The item C_{11} likewise includes two items, the income of unincorporated enterprises, D_{11}, and an inventory calculation adjustment, E_{11}. These inventory valuation adjustments, E_{11} and J_{11}, are put in because in business accounts inventories are valued by methods different from those in the national income accounts, where inventories are valued at weighted average prices.

The next concept is that of personal income, I_{13}. This consists of most of the components of national income: compensation of employees, B_{11}, is divided into wage and salary disbursements, B_{14}, plus an item called other labor income, H_{14}, plus an item not listed in the accounts but which may be called nonpersonal contributions to Social Security, S'_{14}, plus also an item (= 0 in 1964) called excess of wage accruals over disbursements, D_{13}. S'_{14} is equal to contributions for social insurance, C_{13}, minus personal contributions for social insurance, S_{14}, so we have:

$$S'_{14} = C_{13} - S_{14} \tag{6}$$

$$B_{14} + H_{14} + S'_{14} + D_{13} = B_{11} \tag{7}$$

Then to $B_{14} + H_{14}$ we add C_{11} (business and professional income, adjusted), F_{11} (income of farm proprietors) and G_{11} (rental income of persons). Then we add an item called personal interest income, M_{14}, which consists of the item K_{11} of the national income breakdown (net interest), plus F_{13} (net interest paid by government), so that

$$M_{14} = K_{11} + F_{13} \tag{8}$$

Then we add a further item, called total transfer payments, N_{14}, which is equal to government transfer payments to persons, E_{13}, plus business transfer payments, J_{12}, so that

$$N_{14} = E_{13} + J_{12} \tag{9}$$

From this we subtract S_{14} (personal contributions for social insurance) and we add not H_{11} (corporate profits adjusted), but only G_{13} (dividends). In lines 6 and 7 of Fig. 6, note that as we have included S'_{14} in line 6, we must deduct it in line 7. In line 6 we have divided N_{14} into its components, $K_{11} + F_{13}$. We then have

$$I_{13} = B_{14} + H_{14} + C_{11} + F_{11} + G_{11} + G_{13} + M_{14} + N_{14} - S_{14} \tag{10}$$

By combining equation (10) with equations (4)–(9), we get the relation between national income and personal income:

$$I_{13} = A_{11} - H_{11} + G_{13} + F_{13} + (E_{13} + J_{12}) - C_{13} - D_{13} \tag{11}$$

We can see this in Fig. 6. Start from the right-hand boundary of national income (lines 3–5), subtract H_{11} (line 5) and add G_{13}. Then in line 6, add F_{13} and N_{14} ($E_{13} + J_{12}$) and subtract C_{13} in line 7 to get the left-hand boundary of personal income.

Disposable personal income, A_7, is defined simply as personal income less personal taxes (line 8), or

$$I_{13} = A_7 + B_{15} \tag{12}$$

and personal saving (C_7) is defined as disposable personal income less personal consumption expenditure, or

$$C_7 = A_7 - B_1 \tag{12a}$$

This is shown in lines 8 and 9 of Fig. 6.

The Sources and Uses of Savings Identity

From Fig. 6 it is not too difficult to demonstrate the "sources and uses of savings" identity for the United States. Let us start with a new concept, net nondisposable income, N_i (line 9), defined as the excess of net national product over disposable personal income, or

$$N_i = E_{12} - A_7 \tag{13}$$

Looking now at the derivation of the various concepts in lines 3–9, we see that

$$N_i = G_{12} - F_{12} + J_{12} + T_7 + H_{65} + K_{65} + J_{11} - F_{13} - E_{13} - J_{12} + D_{13} + C_{13} + B_{15} \tag{14}$$

We also have from equations (3) and (4)

$$R_7 = M_7 - N_7 - J_1 \tag{15}$$

Then from equation (10) and from the equation for total government receipts, M_7

$$M_7 = G_{12} + C_{13} + B_{15} + H_{65} \tag{16}$$

we have

$$R_7 = G_{12} + C_{13} + B_{15} + H_{65} - E_{13} - G_7 - F_{13} - F_{12} - J_1 \tag{17}$$

Let us now define net business savings, D_{7n}, as gross business retained earnings, D_7, less the capital consumption allowance, B_{12}:

$$D_{7n} = D_7 - B_{12} \tag{18}$$

Then

$$D_{7n} = K_{65} + J_{11} + D_{13} \tag{19}$$

Now combining equation (13) with (17) and (19) we have

$$N_i = R_7 + G_7 + D_{7n} + J_1 + T_7 \tag{20}$$

This we see in line 9 of Fig. 6. R_7 is negative (− 3.9 billion), T_7 is likewise negative (− 1.8 billion). Now we define total net investment, H_{18n}, as total gross investment, H_{18}, minus capital consumption allowance, B_{12}, and total net private saving, B_{18n}, as total private saving, B_{18}, less the capital consumption allowance, B_{12}. We then have total net private saving, B_{18n}, as the sum of personal saving, C_7, and net business savings, D_{7n}; net

private domestic investment, C_{1n}, we define as gross private domestic investment, C_1, minus capital consumption allowance, B_{12}; K_7 is defined as foreign net transfers by government, G_7, less net exports, I_1; subtracting K_7 from C_{1n} gives us H_{18}. That is

$$H_{18n} = H_{18} - B_{12} = C_{1n} - K_7 = C_{1n} - G_7 - I_1 \tag{21}$$

$$B_{18n} = B_{18} - B_{12} = C_7 + D_{7n} \tag{22}$$

Then as

$$E_{12} = B_1 + C_{1n} + I_1 + J_1 \tag{23}$$

$$T_7 + R_7 + E_{12} - B_1 = J_1 + G_7 + D_{7n} + C_7 \qquad \text{(line 9)}$$

$$= J_1 + G_7 + B_{18n} \qquad \text{(line 10)}$$

$$= T_7 + R_7 + J_1 + I_1 + C_{1n} \qquad \text{(line 11)}$$

$$= T_7 + R_7 + J_1 + G_7 - K_7 + C_{1n} \qquad \text{(line 11)}$$

$$= T_7 + R_7 + J_1 + G_7 + H_{18n} \qquad \text{(line 11)}$$

We see immediately (lines 10 and 11) that

$$B_{18n} = T_7 + R_7 + H_{18n} \tag{24}$$

That is, net private saving equals net investment, plus the government deficit, plus the statistical discrepancy. If we add the capital consumption allowance to both sides we have the identity in the gross form

$$B_{18} = T_7 + R_7 + H_{18} \tag{25}$$

This is the form in which it is found in the United States national income accounts.

QUESTIONS AND EXERCISES

1. Construct identities and a diagram like Fig. 6, relating the concepts of gross disposable income A'_7 (p. 34), gross saving, B'_{18} and total investment, H'_{18} to the concepts of Fig. 6.
2. Class exercise: List the GNP and all its components mentioned in this chapter on a duplicated work sheet. Let each member of the class take a different year and fill in the numerical values, using the *Economic Report of the President,* and then list and check the various identities relating these variables.
3. Discuss the effect, if any, on the various items of the national income accounts of the following events:

   a. A corporation makes \$1000 worth of shirts with a labor cost of \$300, a materials cost of \$500, and a profit of \$200; \$800 worth are sold at home, \$100 worth abroad.

   b. The government spends \$10 million in building a dam; \$3 million is

spent for labor, $6 million for materials, $1 million for land.

c. An exporter buys $10,000 worth of cloth from a manufacturer and ships it abroad.

d. The government increases its tax collections by $1 million and increases its expenditures by the same amount. (Note: Where further assumptions are necessary to obtain specific answers, make these assumptions explicit.)

4. What is the significance of the various national income concepts (GNP, NNP, etc.) for employment theory?
5. In calculating the amount of GNP originating in the banking system it is necessary to debit depositors with the actual cost of maintaining their account, rather than with the bank's service charges. Why?

CHAPTER 3

MACROECONOMIC MODELS: MODELS OF PRODUCTION AND CONSUMPTION

The previous two chapters have been devoted essentially to the problems involved in the classification, definition, and measurement of economic aggregates. Mere classification, however, is not enough. We want to know not only "what" but "why" and "how." It is not enough to have a good system of definitions, even a good system of accounts and measurements. We also want to know something about the casual factors involved in economic systems. It is not enough to measure the national income and its distribution, for instance. We are also interested in why the national income and its various components are what they are, and why they exhibit their various changes and fluctuations. The construction of economic models is an attempt to give some answers to the question of the determinants of the economic system—some answer, that is, to the "why and how" questions.

MACROECONOMIC MODELS

What, then, do we mean when we ask, Why is the national income what it is? We may mean many things, depending on the level of argument. In economic analysis, however, we feel that we have a satisfactory answer to the above question if we can show that there is a system of relationships among the various variables of the system *which are satisfied* if these variables are what they are. Such a system of relationships is called a "*model.*" There are two main kinds of these models: equilibrium models and dynamic models. In equilibrium models the system of relationships can be expressed as a set of equations or identities which can

only be satisfied by one set (or at most a limited number) of values of the various variables which the equations relate. In order to obtain these conditions there must be an equal number of equations and unknowns, or variables, to be determined. In dynamic models the values of present variables are related in definite ways to values of these and other variables in the past, which in turn are related to variables in the more remote past, and so on in an infinite regression. The next two chapters will deal mainly with equilibrium models; dynamic models are discussed in Chapter 8.

Identities

The concept of a model will become clear as we proceed to examples. The relationships which constitute a model are of two kinds, *identities* and *behavior equations.* All models seem to possess them both. Identities are true by definition. They are derived generally by taking some aggregate and expressing it as a sum of parts, or as a product of components. If then we put a whole aggregate equal to the sum of its parts, or if we equate two different breakdowns of the same aggregate, we shall have an identity. Thus any human population (N) can be divided into people called Smith (S) and people not called Smith (T), and we can write[1]

$$N \equiv S + T \tag{1}$$

The same population can also be divided into people called Jones (J) and people not called Jones (K), so that we have

$$N \equiv J + K \tag{2}$$

from which we might derive other identities, such as

$$S + T \equiv J + K, \quad \text{or} \quad S - J \equiv K - T \tag{3}$$

We seem even to have derived a theorem, that the excess of Smiths over Joneses in any population must be equal to the excess of non-Joneses over non-Smiths. These particular identities are not very interesting, though nobody can deny their truth. They are not interesting because the aggregates which they relate are not particularly interesting—to the economist, at least. If it were true, for instance, that the number of Smiths and Joneses were casually or empirically related, the above identities might become part of a model and would become interesting.

Suppose, for instance, that the Smiths formed a tight ruling class, and had an inflexible rule (enforced by execution) that the number of non-Smiths should always equal the number of Smiths, and furthermore kept

[1] Identities are generally indicated by the sign $\equiv$, which should be read "is identical with."

their own numbers always at 1000, we would have a "model" from the conditions of which the equilibrium value of all the variables could be deduced. It would be expressed in three relationships: the identity $N \equiv S + T$, and two behavior equations, $S = T$, and $S = 1000$. As there are three unknowns, N, S, and T, these three equations can be solved—that is, values of N, S, and T which satisfy all three relationships simultaneously can be discovered. In this case there is only one solution: $N = 2000$, $S = 1000$, $T = 1000$.

Behavior Equations

Behavior equations, then, are relationships which express certain empirical facts about the behavior of the individuals of the system. They are not *necessarily* true, as identities must be; they derive their validity from the empirical observation that if they are *not* satisfied, behavior, and therefore change, will take place which will not cease until they *are* satisfied. Thus suppose in the above slightly absurd example that the total population was 2500, with 1000 Smiths and 1500 non-Smiths. Following their rule of behavior the Smiths would have to kill off 500 non-Smiths in order to bring their numbers to equality, which would reduce the population to 2000. If the population were 2500, divided between 1250 Smiths and 1250 non-Smiths, both rules would come into play: 250 Smiths would be killed off to reduce the Smith population to the legal 1000, and 250 non-Smiths killed off to bring the numbers of the two groups to equality. If the total population were 1500, say 750 Smiths and 750 non-Smiths, the Smith population would be allowed to grow until it reached 1000, and the non-Smith population would be allowed to grow equally, bringing the total population again to 2000. It is evident that the equilibrium is a stable one. If N is *not* 2000, forces are brought into play to restore that figure. What the population will be on any given date, of course, depends on the dynamic relationships involved—on whether, for instance, the movement toward equilibrium takes place quickly or slowly. In equilibrium models, however, we consider only the equilibrium position itself. The problem of the movements toward (or away from) equilibrium will be discussed in Chapter 8.

Adding New Variables

The models can be made as complicated as we wish by adding new variables and new relationships, as long as we add one independent equation or identity (that is, one that cannot be derived from the existing relationships) for each new variable. Thus suppose we add a new variable to our previous model—say, the number of Joneses, J. In order to make the model determinate we must find a relationship of some kind between

J and the other variables. Suppose for instance that the only function of the Joneses was to cut hair for the Smiths, and that each Jones could handle 20 Smiths. Then in equilibrium we would have an additional relationship, $20\,J = S$, as we know that in equilibrium $S = 1000$, $J = 50$. If now we wish to determine the number of non-Joneses, K, this new variable can be taken care of with the identity, equation (2), so that $K = 1950$.

The above illustration is, of course, farfetched in the extreme, in that the behavior equations postulated are most unlikely to be found in any real society. An absurd case, however, illustrates the generality of a principle better than a realistic one, and it is not difficult now to go on to more realistic models involving economic variables.

The Basic Model of Macroeconomics

The simplest and perhaps the most important macroeconomic model is one of three relationships and three variables. Let P be the total physical output of a society in a given period. Its exact definition and the exact degree of "netness" involved in its definition is not particularly important. Let C be the amount of this physical output consumed during the period, and A be the amount accumulated—that is, added to stocks. Then we have an identity:

$$P \equiv C + A \tag{4}$$

The concepts, of course, must be defined in such a way as to make this identity necessary. What it states in effect is that whatever has been produced in a given period must either have been consumed in that period or must be still around somewhere. Production is conceived, as we have seen, as a gross addition to total stock; consumption as a subtraction from total stock; and accumulation, therefore, is the difference between what has been added and what has been subtracted.

In terms of the concepts of national income accounting we could define P as the net national product, C as the sum of personal consumption expenditures (B_1) and government purchases (J_1), and A as the sum of net private domestic investment (C_1) and the net exports (I_1).

Consumption and Investment Functions

In addition to the identity, equation (4), we need two behavior equations:

$$C = F_c(P) \tag{5}$$

$$\bar{A} = F_a(P) \tag{6}$$

We do not postulate an exact algebraic relationship, but a general function. Equation (5) is read "C is a function of P." It means simply that

the amount of consumption depends on the amount of production in a definite way, so that for each amount of production there is only one amount of consumption which is consistent with it.[2] A behavior equation is not much use, of course, unless there is some assumed behavior behind it. The link here is that decisions to consume on the part of individuals, in so far as they have freedom to choose, are very closely related to their individual incomes. Generally speaking, the larger a person's income, the larger his consumption. Personal incomes, however, are very closely related to total output, P. For any given value of P, therefore, there is likely to be a given value of personal incomes, and for each value of personal incomes there will be a given amount of consumption. It is not unreasonable, then, to assume that with a given distribution of income and stable patterns of individual consuming behavior there is a fairly definite relationship between total output and the decisions which result in consumption, and hence the volume of consumption itself.

The behavior patterns which underlie the accumulation equation, (6), are more complex, and more dubious. What the equation states is that for any given level of output there is only one value of total accumulation which is *consistent with* the given level of output. This assumes that at each level of output there is some level of accumulation which is satisfactory to the people who are holding the additions to stocks of goods—call it $\bar{A}$

We then need an additional equation which expresses the conditions of equilibrium,

$$A = \bar{A} \tag{7}$$

that is, the "actual" level of accumulation must be equal to the satisfactory or equilibrium level.

If the actual level of accumulation is greater than the satisfactory level, the people who hold goods will feel that they are holding too much—they will feel "overstocked." Manufacturers, wholesalers, retailers, and householders may feel that they are holding too large stocks of goods and will try to reduce these stocks. *How* they try to reduce stocks is a critical question in the dynamics of the system. For the moment, however, we will assume that the simplest way to reduce stocks is to cut back production. We have supposed that consumption is given (for each value of P) so that there is not much opportunity to reduce stocks by expanding consumption. Similarly, if the actual level of accumulation is less than the

[2] In strict mathematical terms, the existence of a general function relating two (or more) variables means only that there is a *limited number* of possible combinations of values of the variables. In the present case it is reasonable to assume that there is in fact only one value of consumption corresponding to each value of output. When this limitation is placed on the nature of the function, it is said to be *monotonic*.

satisfactory level, people will feel "understocked" and will try to increase these stocks. We assume for purposes of this model that they can increase stocks only by increased production. Thus in the first case if householders feel overstocked, say with shirts, they reduce purchases of shirts until some of those they have wear out. The reduced purchases result in an overstock on the part of the retailer, who reduces purchases from the wholesaler; reduced purchases from the retailers lead to overstock at the wholesaler, who reduces purchases from the manufacturer. The manufacturer responds to his resulting overstock by laying off men and decreasing the output of shirts. By a similar chain an understock anywhere in the system will lead to increased production, as long as there are unused resources.

We see, therefore, that if the behavior equations are *not* satisfied certain *changes* in behavior will take place, and if the equilibrium is a stable one these changes will operate to adjust the values of all the variables of the system in a direction which makes the equations more nearly satisfied. This process will presumably go on until the equations are satisfied, at which point there will be no further changes in behavior and the system will reproduce itself indefinitely in equilibrium until the equations change.

Arithmetical Example of Macroeconomic Equilibrium

This simple model can be illustrated by arithmetical and graphical examples. Suppose that Table 3 shows the relationship in some society between total output (P) and those levels of consumption (C) and accumulation $(\bar{A})$ which are consistent with each level of output.

TABLE 3. OUTPUT EQUILIBRIUM

1. Output (P)	0	20	40	60	80	100	120	140	160	180	200
2. Consumption (C)	60	73	87	97	106	113	120	126	130	134	137
3. Willing accumulation (investment) $(\bar{A})$	−30	−13	0	7	15	21	24	27	30	32	34
4. Total absorption $(\bar{A} + C)$	30	60	87	104	121	134	144	153	160	166	171
5. Actual accumulation $(P - C)$	−60	−53	−47	−37	−26	−13	0	14	30	46	63
6. Surplus (+) or deficit (−) accumulation $(P - (\bar{A} + C))$ or $((P - C) - \bar{A})$	−30	−40	−47	−44	−41	−34	−24	−13	0	+14	+29

The relationship between output and consumption, $C = F_c(P)$, is known as the *consumption function*. The relationship between output and "willing accumulation," $\bar{A} = F_a\ (P)$, is often called the *investment function*, investment being defined as the amount of voluntary or willing ac-

cumulation which is consistent with a given level of output.[3] Thus the table should read: if the output were 0, the society would be willing and able to consume 60, and would be willing and able to invest −30 (that is, to disinvest, or decumulate 30); if the output were 20, the society would be willing and able to consume 73, and invest −13; and so on. The fourth line of the table shows the algebraic sum of consumption and investment. It is useful to have a name for this quantity, so I have called it "total absorption." If we think of consumption and investment (willing accumulation) as the only two ways in which the product can be absorbed or disposed of willingly, the total of consumption and investment together shows the total amount of product which will be voluntarily absorbed by the system. If the actual output is greater than the total (willing) absorption, there will be *unwanted* accumulations; if actual output is less than total (willing) absorption, there will be unwanted decumulations, and, as we have seen, these will produce changes in behavior which will change output itself.

Assumptions of the Model

In constructing Table 3 we made certain assumptions about the *nature* of the consumption and investment functions. These assumptions are plausible, but not absolutely necessary. The actual shape of the function is a matter of fact in each case which can only be determined by empirical research. Thus we assumed that at zero output, should such a state of affairs be even possible, there would still be some consumption (60), as existing capital depreciates and people must eat even if they are not producing. In a state of such extreme depression, however, it is most unlikely that anyone would be willing to accumulate. Indeed, the reverse is assumed—that at very low levels of output the society is only willing to "decumulate," i.e., to live on its existing stocks. As we advance (hypothetically) to higher levels of output we assume that both consumption and investment rise. In the example we have assumed further that they both rise at a decreasing rate, so that a unit rise in output at high levels of output produces a smaller increase in C and $\grave{A}$ than does a unit rise in output at low levels. This assumption, though plausible, is in no way necessary for the argument, as we shall see later.

Surplus or Deficiency in Accumulation

If we now subtract the total (willing) absorption figure from the output figure we get a measure of "surplus" or "deficiency" in actual accumula-

[3] In the Keynesian system investment is not generally regarded as a function of output, but of other things, such as the rate of interest. These complications will be introduced later.

tion (item 6, Table 3). If there is an accumulation surplus (as there is in the table at levels of output above 160), this means that at these levels of output actual accumulation is greater than the amount of accumulation which the people of the society are willing to accept voluntarily. If there is an accumulation deficit, as there is below an output of 160, then actual accumulation is less than the desired level of accumulation. We can suppose that actual accumulation (item 5) is the difference between output and consumption at each level of output. Then the accumulation surplus or deficit can also be obtained by subtracting the figures of item 3 (willing accumulation) from the corresponding figures of item 5 (actual accumulation).

EQUILIBRIUM OF THE MODEL

We can now see clearly that there is only one output figure at which the system is in equilibrium. At an output of 160 total absorption is also 160; willing accumulation is equal to actual accumulation (30) and there is no surplus or deficit accumulation. At all outputs lower than 160 total absorption exceeds output; the amount that the society is willing to accumulate exceeds what it actually succeeds in accumulating, and there is deficit accumulation. Thus at an output of 140, consumption is 126, so that actual accumulation is 14. The amount that people wish to accumulate, however, is 27. The society is understocked and output will increase. Similarly if output is 180, consumption is 134, so that actual accumulation will be 46. People only wish to accumulate 32, so that there will be a surplus accumulation of 14, and output will be reduced to try to get rid of the surplus.

GRAPHIC ILLUSTRATION. This model can also be illustrated graphically. In Fig. 7 output is measured on the horizontal axis, and consumption and accumulation on the vertical axes. It should be remembered that these measurements are indices of real output or consumption, and that the indices chosen must be consistent. This means that the relative price structure used in computing an index must be the same for all the indices. CC' is then the consumption curve, corresponding to the consumption function of items 1 and 2, Table 3. II' is the investment curve, or willing accumulation curve. BB' is the vertical sum of these two curves, or the total absorption curve (items 1 and 4 of the table). We now draw a line OB_e from the origin at an angle of 45 degrees to each axis, cutting the curve BB' at B_e. The 45 degree line represents the identity $P \equiv C + A$. The point B_e then gives the equilibrium output, OO_e, for at this point output, $OO_e = O_eB_e = O_eC_e + C_eB_e = O_eC_e + O_eA_e = C + A$. The equilibrium point can also be found in two other ways. One method is to draw

the actual accumulation curve AA' (items 1 and 5 of Table 3). This is given by the vertical distance between the line CC' and the 45 degree line OB_e. Where the curve AA' intersects the curve II'—i.e., where actual accumulation is equal to desired accumulation—is the point of equilibrium.

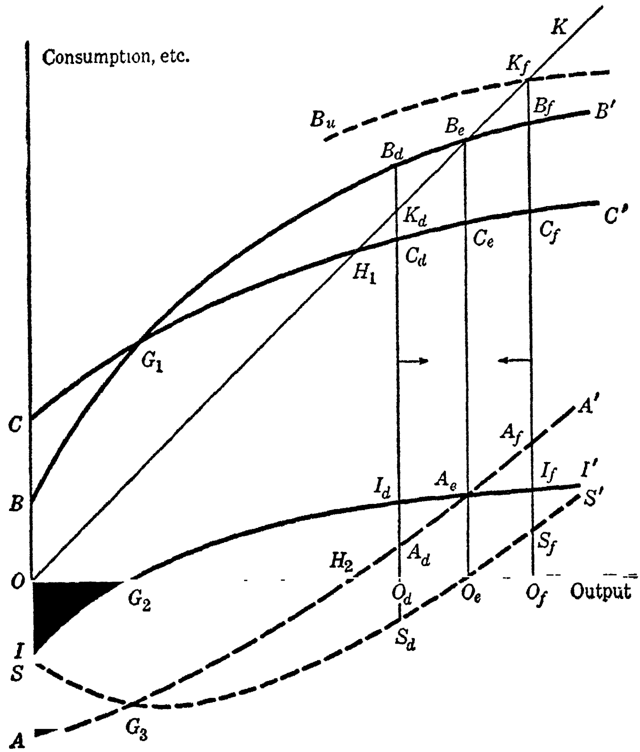


Fig. 7. Output Equilibrium

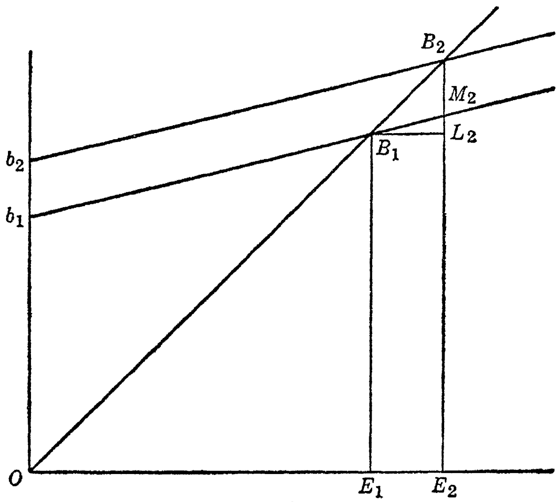
Another method is to draw the surplus-deficit accumulation curve SS' (items 1–6 of the table). Where this cuts the horizontal axis—i.e., where the surplus or deficit accumulation is zero—is the point of equilibrium.

This diagram also illustrates the stability of the equilibrium. Thus suppose output were below the equilibrium output, say at OO_d. Consumption would be O_dC_d. Actual accumulation, therefore, would be C_dK_d, where K_d is on the 45 degrees line, as $O_dK_d = OO_d$. Desired accumulation is $C_dB_d (= O_dI_d)$. There is, therefore, a deficit of actual accumulation below the desired level amounting to B_dK_d $(= I_dA_d = O_dS_d)$. This deficit will cause an increase in output, and the whole system will therefore move toward the equilibrium. Similarly, if output is above the equilib-

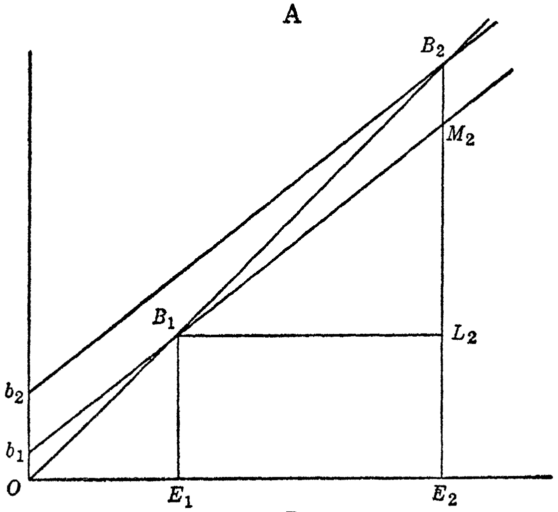
brium, say OO_f, there will be a surplus accumulation ($B_fK_f = I_fA_f = O_fS_f$) which will cause cutbacks in output and again return it to the equilibrium value OO_e.

Effect of a Rise in the Absorption Curve

This model, simple as it is, is so powerful in the explanation of movements in aggregate incomes, even in the explanation of business cycles,



A



B

Fig. 8. The Multiplier

that it is worth examining some of its properties in greater detail. We see first that a rise in either the consumption or the investment curves, or in both, will raise the equilibrium output. Thus in Fig. 8 we suppose a total absorption curve b_1B_1, with the point of equilibrium at B_1, OB_1 being the 45 degree line. If the total absorption curve rises to b_2B_2, that is, if the total of consumption and willing accumulation is larger, *at each level of output,* than it was before, the equilibrium output will rise from OE_1 to OE_2. Similarly, a fall in the absorption curve results in a fall in the equilibrium output. These changes in the total absorption curve may have several origins. There may be a spontaneous rise in the consumption function itself—that is, the people of the society may decide to consume more at each level of income than they did before. There may likewise be a spontaneous increase in the investment function—that is, people may become willing to accumulate more goods at each level of income than they did before. There may also, as we shall see later, be an increase in absorption by government without an offsetting decline in absorption by private persons.

The Propensity to Absorb and the Multiplier

Whatever the causes, there is an interesting relationship between the *extent* of the increase in the willing absorption of product and the increase in the equilibrium output which is called forth. In Fig. 8 we have supposed that the total absorption curve is a straight line. Modifying the Keynesian terminology somewhat, we may call the slope of this line the *propensity to absorb.* The propensity to absorb, then, is the increase in total willing absorption of output which resuts from a *unit* increase in output. Thus if an increase in output of 10 units resulted in an increase in willing absorption in consumption and investment of 8 units, the propensity to absorb in this range would be $\frac{8}{10}$, or 0.8. In the figure, if a perpendicular is dropped from B_1 to E_2B_2 at L_2, the propensity to absorb, α, is $\frac{L_2M_2}{B_1L_2}$. The ratio of the vertical increase in the absorption function to the resultant increase in the equilibrium output is called the multiplier. The historic reason for this term is that it was first used in connection with the increase in income (output) which would result from a unit increase in *investment,* investment being assumed to be independent of output, so that the multiplier is the amount by which the initial increase of investment multiplies itself in producing income. In the figure, then, the multiplier, m, is the ratio $\frac{B_1L_2}{M_2B_2}$; that is, it is the amount by which output increases for a unit vertical rise in the absorption function.

There is a relatively simple mathematical relationship between the multiplier, m, and the propensity to absorb, α. Thus we have in the figure $B_1L_2 = L_2B_2$, by construction. It follows then that

$$m = \frac{B_1L_2}{M_2B_2} = \frac{L_2B_2}{M_2B_2} = \frac{L_2B_2}{L_2B_2 - L_2M_2} = \frac{1}{1 - \dfrac{L_2M_2}{L_2B_2}} = \frac{1}{1 - \alpha} \tag{8}$$

It follows immediately from this formula that the closer α lies to 1, the larger will be the multiplier. If $\alpha = 1$ the multiplier is infinite; the absorption curve is then at an angle of 45 degrees and coincides with the 45 degree identity line. In this limiting case the equilibrium itself, of course, is indeterminate. If the propensity to absorb at the equilibrium point is greater than 1, the equilibrium is unstable.

The relationship between the propensity to absorb and the multiplier can be visualized graphically by comparing Fig. 8A, where the propensity to absorb is small (about 0.25) and where the multiplier is also small (about 1.33), with Fig. 8B, where the propensity to absorb is large (about 0.8) and the multiplier, therefore, is also large (about 5). The same upward shift in the absorption curve in both figures, b_1b_2, produces a much larger change in output (B_1L_2) in Fig. 8B than it does in 8A.

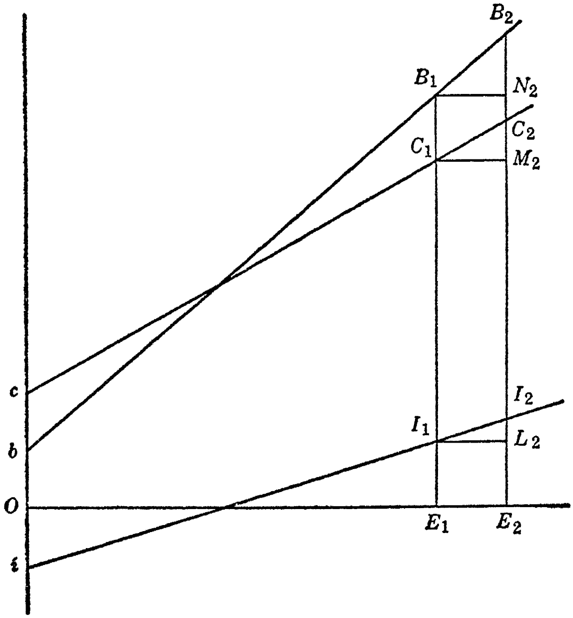


Fig. 9. The Propensity to Consume and to Invest

THE PROPENSITIES TO CONSUME AND INVEST. The propensity to absorb is the sum of two similar quantities—the propensity to consume and the propensity to invest. This is shown in Fig. 9. I_1I_2, C_1C_2, and B_1B_2 are segments (assumed to be straight lines) of an investment curve, a consumption curve, and the derived total absorption curve. From the method of construction we know that

$$E_1I_1 + E_1C_1 = E_1B_1 \qquad \text{and} \qquad E_2I_2 + E_2C_2 = E_2B_2$$

It follows that

$$E_2I_2 - E_1I_1 + E_2C_2 - E_1C_1 = E_2B_2 - E_1B_1, \quad \text{or} \quad L_2I_2 + M_2C_2 = N_2B_2.$$

Dividing through by E_1E_2 or its equivalent we have

$$\frac{L_2I_2}{I_1L_2} + \frac{M_2C_2}{C_1M_2} = \frac{N_2B_2}{B_1N_2}.$$

$\dfrac{L_2I_2}{I_1L_2}$ is the propensity to invest, α_i—that is, the increase in willing accumulation for a unit increase in output. $\dfrac{M_2C_2}{C_1M_2}$ is the propensity to consume, α_c—the increase in consumption per unit increase in output. $\dfrac{N_2B_2}{B_1N_2}$ is the propensity to absorb, α. We have therefore $\alpha = \alpha_i + \alpha_c$. In the Keynesian system proper the propensity to invest is generally taken to be zero unless there is induced investment, of which the propensity to invest is a measure. If the propensity to invest is zero we have $\alpha = \alpha_c$, and

$$m = \frac{1}{1 - \alpha} \tag{9}$$

In fact it seems probable that the propensity to invest is at least a small positive figure, and under some circumstances it may be quite important. The larger it is, of course, the larger also will be the propensity to absorb, and hence the multiplier.

The "Shiftability" of Equilibrium

The multiplier concept was developed, as noted earlier, in the attempt to determine the increase in national output, income, or employment which would result from a given spontaneous increase in investment or in government expenditure. The limitations which have to be placed on the model in order to make it meaningful are so restrictive that in its original meaning it does not seem to be particularly useful. Thus we would have to assume that the consumption function was unchanged in the face of investment changes, and that there was no induced invest-

ment before the formula (9) takes on much meaning. Nevertheless, the multiplier, or what is practically the same thing in a different measure, the propensity to absorb, is an extremely important property of the model, for it measures the degree of shiftability of the equilibrium position. If the multiplier is small, as in Fig. 8A, a given shift in the underlying determinants of the system, as reflected in a change in the position of the absorption curve, will not produce a very large change in the position of equilibrium itself. By contrast, when the multiplier or the propensity to absorb are large, as in Fig. 8B, quite small disturbances in the underlying conditions will produce large changes in the national income or output.[4]

This fact is of great importance for highly industrialized modern economies like that of the United States, as it is a possible explanation of the great instability of these economies. The propensity to absorb in America, for instance, seems to be at least of the order of 0.8. Quite slight changes in the underlying behavior patterns, therefore, may give rise to very large changes in the position of equilibrium; a slight fall in the level of the absorption function may give rise to devastating depression and unemployment, and an equally slight rise may result in full employment, prosperity, and even inflation.

APPLICATIONS OF THE BASIC MODEL

Perhaps the most important application of the basic model is in giving at least a rough explanation of the phenomenon of depression and unemployment. There is nothing in the equilibrium of the model as we have described it which says that the equilibrium position of output is at capacity output. The equilibrium may, therefore, be at a level which is below what the society is capable of reaching, even with the resources which it has available at the moment. If the equilibrium output is below

[4] If the functions are linear, we can write

$$C = C_o + a_c P$$
$$\bar{A} = A_o + a_i P$$

Then as in equilibrium $\bar{A} = A$, and $P = C + A$, we have

$$C + A = C_o + A_o + (a_c + a_i)P = P, \text{ or}$$

$$P = \frac{C_o + A_o}{1 - (a_i + a_c)} \tag{10}$$

C_o and A_o are important parameters which measure the height of consumption or willing accumulation. We see that the higher is consumption and willing accumulation, and the larger the propensities to consume and invest, the larger the equilibrium value of P.

capacity, then, some resources will be unemployed. A measure of unemployment is the difference between capacity output and the actual or equilibrium output. Thus, returning to Fig. 7, page 54, suppose that OO_f represents capacity output. Equilibrium output is OO_e. The difference, O_eO_f represents that output which could be produced but which its not being produced because certain resources are unemployed. It is therefore a measure of the amount of unemployed resources in the system.

Unemployment Equilibrium

If now we inquire *why* there is unemployment, the answer is to be found in a study of what would happen if in fact the system were operating at capacity, i.e., at OO_f. At this level of output we see that consumption and investment (willing accumulation) are insufficient to absorb the whole product. In other words, there is a deficiency in "effective demand." This reflects itself, as we have seen, in unwanted accumulations amounting to B_fK_f (O_fS_f). These unwanted accumulations produce a downward pressure on output, so the economy cannot remain at full employment. At full employment it will be accumulating stocks of goods at a rate faster than its people are willing to take. We see unemployment, therefore, as an attempt on the part of the system to cut down its rate of accumulation to a level that its people are willing to absorb. Output has to be cut back by *more* than the excess accumulation, because cutting back output itself reduces consumption, and the reduction in consumption to some extent offsets the decline in production. Because, however, consumption does not decline as much as production, the reduction in production does bring about a decline in the surplus accumulation, and will go on until this surplus is reduced to zero. If the propensity to absorb is high, consumption and investment decline sharply as output declines, which reduces the effect of the decline in production on accumulation; the decline may therefore have to go a long way before the surplus accumulation is eliminated. Suppose, for instance, that with an output of 200, willing absorption is 180, leaving a surplus accumulation of 20. If now the society cuts back its production to 180, it will still have a surplus accumulation, because willing absorption will also have declined, say, to 165, still leaving a surplus of 15, and production must be cut back further.

This phenomenon is sometimes known as a vicious spiral: the attempt to cut back unwanted accumulation by reducing output has the effect of reducing incomes, which in turn lowers consumption. This fall in consumption in part offsets the fall in output, so output must fall further, and so on, until declining output overtakes declining consumption and even the declining willingness to accumulate.

A Full-Employment Model

A very important question must now be raised regarding the relevance of this model to the actual world. Underemployment equilibrium can only exist if the consumption and investment functions are stable and, in particular, are unaffected by the level of unemployment. It is not difficult to expand the basic model by adding another variable, let us say, unused capacity, U, and then postulating that both consumption and investment are themselves related to unused capacity as well as to output. Let us suppose, for instance, that as long as there is any unused capacity ($U > O$) the consumption plus investment function will rise—that is, in the linear case, $\frac{d(C_o + A_o)}{dt} > O$. Then as these functions rise P will increase, U will decline, and this will go on until $U = O$, when the rise in C_o, A_o, and P will stop. In terms of the graphic analysis this means that in Fig. 7, page 54, for instance, if there is unused capacity, the total absorption curve will move upward until it reaches a position such as B_uK_f, where it intersects the 45 degree line at K_f. The surplus accumulation is reduced to zero in this case not by a decline in output, but by a rise in the whole level of the absorption curve.

The Keynesian Controversy

Much of the dispute between the Keynesian and the classical school of employment theorists (as represented, say, by Pigou) rests on the question of the extent to which there is anything in reality to correspond to the behavior equations outlined above. The full discussion of the behavior patterns involved must wait until we have discussed the operations of the monetary and price systems, but at least an outline can be given here. The Keynesian model assumes that there is no "natural" tendency in the system tending to raise the level of the absorption curve as output falls below capacity, and that if full employment is to be maintained there must be active intervention on the part of government to raise the level of the absorption curve to the point where capacity output will not result in surplus accumulations. We can, of course, if we like, regard this governmental intervention as itself part of an equilibrium system, and assume that a rise in U above a certain critical level will call forth expansionary activity on the part of government, and vice versa. In that case $U = O_e$ would be a true equilibrium, though the behavior equations which determine it rest on political rather than on strictly economic behavior.

The Classical System: Deflation of Prices Rather Than of Outputs

The classical assumption regarding the behavior which underlies the equations is that the development of surplus accumulation, or overstocking, will produce a downward pressure on *prices* rather than on outputs. If all markets are competitive and therefore all prices flexible, this assumption is not implausible. Unwanted stocks of commodities produce deflations in the prices of these commodities, as the owners of these unwanted stocks try to get rid of them by selling them. Similarly, if labor markets were competitive, unemployment would produce a downward pressure on *money* wages (not necessarily on real wages) which would supplement and encourage the general price decline. Thus we see the major consequence of a surplus accumulation, on these assumptions, is a general and widespread decline in the whole price-wage level.

Effects of Deflation

ON CONSUMPTION. The economic consequences of general *deflation,* which is the name given to a general decline in the level of money prices and wages, are complex in the extreme. Nevertheless two major sets of consequences may be expected to affect the consumption and investment functions. In the first place deflation produces a marked redistribution of income and of wealth among different groups and classes of people in the society. People who have incomes which are fixed, or which are relatively inflexible in money terms (i.e., in "dollars"), gain at the expense of those whose dollar incomes are flexible. Thus there are gains to bondholders, debt holders in general, pensioners, people living on annuities, and people with salaries and wages which are set by contract or by convention and changed only under severe pressures or at infrequent intervals. These gains are obtained at the expense of businessmen, farmers, people who live from profits in general, people whose incomes depend on flexible prices or wages. These shifts in the distribution of income on the whole are likely to raise the consumption function, as the people who gain are for the most part the "low savers"—people who consume nearly as much as or more than they produce. The people who lose are on the whole the "high savers," businessmen and farmers, people in the prime of life who are building up an estate and who therefore consume much less than they produce. Deflation, therefore, is likely to shift income from the high savers to the low savers, and therefore will encourage consumption by the people who gain more than it will discourage consumption by the people who lose. The net result therefore is likely to be a rise in the amount that will be consumed at each level of income—i.e., a rise in the consumption curve.

ON INVESTMENT. The effects of the redistribution of income on the investment curve are perhaps more difficult to estimate, but it is probable that the effects will be *unfavorable* to investment. On the whole it is the businessmen and farmers who make the major investment decisions and whose willingness to hold increased stocks of goods, new buildings, and so on, is the main factor influencing the level of the investment function. It is precisely these people who are injured by deflation. The profit maker obtains his income by selling dearer than he buys, usually after some interval of time. If during the interval between purchase and sale all prices have fallen, the chance that the value of sales exceeds the value of purchases is so much the less. In extreme deflation, profits may even disappear altogether, and become losses. Under these circumstances it is extremely unlikely that there will be much desire to increase holdings of goods, for it is precisely the holding of these goods, which are declining in price, that is the source of loss. We may reasonably expect, therefore, that deflation will have an adverse effect, at least in the short run, on the investment function. The net effect of the redistribution of income and wealth effected by deflation, then, depends on the relative magnitudes of the effects on consumption and on investment, which are likely to be in opposite directions and to serve to offset each other. If the rise in consumption is larger than the fall in investment at each level of output, the net effect will be a rise in the total absorption curve, and hence a rise in output toward full employment. If, however, the fall in investment at each level of output exceeds the rise in consumption, the system is in a sad plight. The result of surplus accumulation will be a *decline* in the total absorption curve, an *increase* in the surplus, and a further decline in output. Output will decline in this case until a "floor" is reached, imposed by the necessity for maintaining at least a subsistence level of consumption, and the impossibility of disinvesting at more than the rate set by physical decay. In terms of Fig. 7, if the effect of deflation is to lower investment by more than it raises consumption, instead of the total absorption curve moving upward from B_dB_f to B_uK_f under the impact of unused capacity, the curve will move downward, thus increasing the unused capacity, and will continue to move downward until the favorable effect on consumption just outweighs the unfavorable effect on investment.

The Liquidity Effect

The distribution effect of deflation is not the only one which may be relevant to the absorption function. There is also a "liquidity effect." If the quantity of money is constant, the fall in the price-wage level increases the purchasing power of the existing stocks of money. People,

therefore, find themselves more liquid than before, with a larger proportion of money in their total asset structure, as the money value of other assets has declined. This increased liquidity should encourage both consumption and investment. It is reasonable to suppose that increased money stocks with constant prices will encourage consumption and investment, as the increased stocks of money give people a greater feeling of power to consume and invest. A fall in prices with constant money stocks should have the same result. The liquidity effect therefore may reasonably be supposed to augment the favorable effect on consumption and to offset, at least in part, the unfavorable effect on investment of the distribution effect. It increases the probability that the existence of unused resources will tend to raise rather than lower the total absorption curve, and therefore makes a full employment equilibrium more plausible. We shall return to this topic in a later chapter, when we have developed some further analytical tools. It is sufficient to notice now that the difference between the two systems rests mainly on whether the development of surplus stocks will produce cutbacks in *output* or in *prices*. This is a question of fact, and the facts may differ from place to place and time to time, so that the Keynesian model may be relevant in one society and the classical model in another.

Where labor and commodity markets are highly organised, where collective bargaining is the rule, and where workers offer strong resistance to reductions in money wages, the Keynesian model is more plausible; where all markets are competitive and money prices and wages highly flexible, the classical model may be closer to reality. In most modern societies there has been a strong tendency for money prices and wages to become more rigid, especially in resisting downward movements, and it is this more than anything which has made the Keynesian solution more attractive.

Inflation in the Basic Model

The basic model throws some light also on the problem of inflation, though here again a full treatment must be delayed until further tools have been developed. Inflation results when the equilibrium output of the basic model is *greater* than capacity output. Obviously output itself cannot expand beyond capacity, and hence if at capacity output there is still a deficit in accumulation, and the society is still wishing to consume and accumulate more than it is in fact producing, changes must take place in the society which will tend to get rid of the deficit accumulation in other ways than by an expansion of output. Thus suppose in Fig. 7, page 54, again that capacity output were OO_d. At this output there is a deficit in accumulation of K_dB_d, indicating that

people would like to accumulate more than they are succeeding in doing, and there is therefore a deficiency or shortage in stocks. In this case, of course, the output adjustments are manifestly impossible, except in so far as it is possible, especially under conditions of stress such as a war, for a society to expand its capacity beyond what it formerly believed it to be. If there is still a deficit in accumulation where the limits even of extreme capacity have been reached, the result will be either inflation or government controls designed to lower the consumption and accumulation curves.

The Mechanism of Inflation

The mechanism of inflation will be described in more detail in Chapter 4. Its impact on the consumption and investment functions is all that concerns us here. This impact, as in the case of deflation, is likely to operate through two main mechanisms—the redistribution of income on the one hand, and a liquidity effect on the other. Both operate in an opposite direction to the way they operate in deflation. Thus inflation redistributes income toward the profit makers and the holders of goods, and away from the *rentiers*—the debt holders, pensioners, annuitants, and people with relatively fixed money incomes. In this case, therefore, the gainers are likely to be the high savers and the losers the low savers. The effect of the redistribution, therefore, is likely to be a decline in consumption at each level of output. The effect on investment, however, is likely to be in the opposite direction: the high profits of inflation encourage investment; the holding of goods becomes profitable and so the willingness to hold *more* goods—i.e., to accumulate—increases.

Again the net effect of the redistribution depends on the magnitude of these changes: if the effect on consumption is large and on investment is small, the net effect will be a decline in the total absorption curve and the inflation will therefore be effective in lowering the deficit accumulation and in moving toward an equilibrium. If, however, the effect on consumption is small and on investment is large, the inflation adds fuel to its own flames—the more it proceeds, the greater the deficit in accumulation and the greater the inflationary pressure. In such a case an inflation may lapse into "hyperinflation," in which the rate of rise in the price level approaches a limit set by physical circumstances. This limit, if we may judge from the experience of hyperinflations, seems to be about 100 percent per week.

Liquidity Effect in Inflation

Inflation also may have a liquidity effect. If the rise in prices proceeds faster than the increase in the money stock (which is frequently

the case), then even if the money stock is increasing people find themselves "less liquid" because the purchasing power of their money stock declines. This may make them less willing to consume and to invest at each level of income, as they feel their power of consuming and investing is less. The liquidity effect, therefore, operates to lower both the consumption and investment functions, and hence has a doubly restraining effect on total absorption. It is an effect which lessens the deficit accumulation and helps to restore equilibrium. We may find, therefore, that there is some constant rate of inflation at which the total absorption curve is pulled down to the point where there is no deficit or surplus in accumulation.

QUESTIONS AND EXERCISES

1. Construct identities relating the following sets of variables:
   a. The average yield of wheat per acre, y, the total number of acres harvested, A, and the total production of wheat, W.
   b. The average number of vehicles per mile of highway, n, the average speed of the vehicles, v, and the density of traffic (number passing a given point per hour), D.
2. Express the variables of a balance sheet as an identity. Can the balance sheet identities of various individuals be summed to produce a consolidated balance sheet identity? If so, what variables will it contain?
3. What is the interpretation of the points of intersection G_1, G_2, G_3 and H_1, H_2 in Fig. 7, page 54.
4. In Table 3, page 51, suppose that government is able to absorb 14 units of product from the economy without causing any decline in consumption or investment. Calculate and tabulate the total absorption, the actual accumulation, and the surplus or deficit accumulation at each level of output. What now will be the equilibrium output?
5. In the model of Table 3, suppose that government absorbs 20 units of output, but that this absorption is financied in such a way that it causes a drop of 7 units of consumption and 3 units of investment at each level of output. Using the graphical technique of Fig. 7, find the equilibrium output.
6. Suppose that the total absorption curve BB' of Fig 7 was moved vertically downward until it was tangent to the line OK. Discuss the nature of the equilibrium at the point of tangency. How would you interpret the model if the curve BB' lay wholly below the line OK?
7. Suppose that a rate of inflation of x percent per annum will lower the total absorption at each level of output by $\frac{x\%}{2}$. Suppose now that in the conditions of the model of Fig. 7 and Table 3, the capacity output of the system was 140 units. What rate of inflation would be necessary in order to equate total absorption with output? Draw the family of curves of Fig. 7 under these new conditions.

8. Suppose we have consumption and investment functions as in footnote 4.

$$C = 100 + (0.7)P$$
$$\bar{A} = -20 + (0.1)P$$

Calculate (1) the equilibrium value of P
(2) the investment multiplier

Suppose now that capacity output is $P_c = 500$. Assuming no change in the marginal propensities, how much will $C_o + A_o$ have to rise in order to achieve full employment? Suppose that A_o does not change, that C_{t+1} and C_t are values of C_o in successive time periods, and that $C_{t+1} - C_t = (.5)\ (P_c - P)$. Calculate the equilibrium value of P for a number of successive periods and show that it converges on full employment.

CHAPTER 4

MACROECONOMIC MODELS: MODELS OF MONETARY CIRCULATION AND EXCHANGE

The Exchange System

The economic models of the previous chapter were *nonmonetary* in the sense that all the variables in them were aggregates of goods. Production, consumption, investment, were defined as aggregates of goods created, destroyed, or accumulated. Income was identified with output as an aggregate of real goods produced.[1] In expanding the model we must now turn to the aspects of the system which are involved in the circulation of money and the exchange of goods. The basic model of Chapter 3 makes no explicit mention of exchange, which is the process whereby the *ownership* of goods and of money and other assets is transferred and circulated among the various individuals of the society. In exchange no physical, identifiable asset is created or destroyed, even though utilities are generally created, but the previously produced assets change owners in the course of their life history from their birth in production to their death in consumption.

THE NATURE OF MONEY

Money as a Measure of Value

In this process of exchange, money plays an essential role in any developed economy. Barter—that is, the exchange of goods for goods,

[1] It is curious that what is really the basic model of the Keynesian economics should be essentially nonmonetary, when even Keynes himself, at least for a long period in the development of his thought, regarded his main contribution as the development of a "monetary" economics as opposed to the "real-goods" economics of the classical school. In fact, however, the monetary elements in the Keynesian system are essentially subordinate to the "real" elements.

or of any nonmoney asset for any other—is important in primitive societies and is never wholly absent even in the most advanced economies. A pure barter system, however, is rare even among primitive societies. In the course of the development of economic institutions the evolution of a standard and measure of value takes place at a very early date. A pure barter economy only seems to be possible where exchanges are rare and the commodities exchanged are highly diverse. As soon as anything like a market develops—that is, frequent and repeated exchanges of more or less standardized commodities—the principle of arbitrage will operate (see Volume I, Chapter 8) to bring the exchange ratios of various commodities into rough consistency with each other. When this happens it becomes possible to use one commodity as a measure of value, and a convention arises whereby prices are quoted in terms of the standard commodity, even though this may not participate in all exchanges. A great variety of such standard commodities have been used in different societies—cattle ("pecuniary" comes from the latin word *Pecus,* cow), seashells, beads, stones, tobacco, cloth, grain, and all the common metals.

All Measurement as Valuation

The same process takes place in the development of almost any kind of measurement. Measurement is the expression of equivalence in some quality between different objects. Thus we might say that a house was "10 pianos long" or "17 sheep long" or "800 thimbles long," meaning that we had established a ratio of equivalence in length between the house and 10 pianos, or 17 sheep or 800 thimbles. It is very convenient, however, to have a single measure, and in the case of length some common object, such as a foot or a hand, generally first serves this purpose. Similarly in the case of value we might say that the house was worth 5000 yards of cloth or 10,000 bushels of wheat or 25 horses, expressing here an equivalence in exchange. In this case also it is convenient to have a single measure, and a society soon comes to adopt a convention to use one commodity to measure all values. If then we know the equivalence of two objects in terms of the measure, we know their equivalence in terms of one another. If a house is 50 feet and a table is 5 feet long, we know immediately that the house is as long as 10 tables. Similarly, if a house is worth 30 cows and a horse is worth 2 cows, we know that a house is worth 15 horses.

The Abstract, Standard Measure

As measurement is essentially an abstract process and is not necessarily tied to any physical object as a measure, it is not surprising to find in the course of development that concrete measures give way to

abstract units of measurement. This is mainly because all concrete measures are apt themselves to vary in the quality measured. If we measure lengths by actual feet or paces, we run up against the problem that one person's foot or pace is different from another's. Consequently the idea of a "standard" foot emerges, which may not be the length of any actual foot. Similarly an abstract measure of value emerges in the shape of a monetary unit, the "dollar" or "pound" or "franc." The "dollar" as a unit of measurement is not a commodity, not even a certain weight of gold. It is essentially an abstract unit by which the values, or exchange equivalences, of assets are compared.

Money as a Medium of Exchange

It is quite possible for a society to have a unit of account, or a measure of value, which is purely abstract and is not represented in any physical form. In England, for instance, the price of luxury articles is often quoted in guineas, although the guinea coin (worth 21 shillings) disappeared many generations ago. Usually, however, the unit of account is embodied in some physical object or asset, in which case it is said to be a *medium of exchange* or a *store of value*. In societies where a commodity is used as a unit of account, such as gold or tobacco, the same commodity in its physical form also generally becomes a medium of exchange, that is, an *intermediary* between final exchanges. The worker exchanges his labor for the money, whatever it is, not because he wants it for its own sake, but because he can buy other things with it. His final exchange is labor for goods. This, however, is broken down into two sets of exchanges, labor for money and money for goods. The medium of exchange in a society, then, is that asset or group of assets which is wanted primarily because it is readily exchangeable. This attribute of ready exchangeability is called "liquidity."

Liquidity

It is important to realize that liquidity is a *quality* of assets which they may possess in greater or less degree. Liquidity is not a very clear or easily measurable concept; yet it is of the utmost importance in understanding the nature of money.

Consider, for instance, the difference between an automobile worth $2000 and a roll of twenty hundred dollar bills. Both objects are worth $2000. What this means is that each of these objects could be exchanged directly or indirectly for $2000 worth of any other commodity or group of commodities which are offered on the market. The amount in both cases represents a "claim" on the other goods and services of society. In both cases there is an understanding, implicit in the whole system of

exchange relationships and valuations, that I can find someone who will give me $2000 worth of something else in return for what I have. The difference, however, lies in the fact that the twenty hundred dollar bills could quickly and easily be exchanged for anything that is offered for sale in the United States, whereas the automobile could not easily be exchanged for anything except money. That is to say, the twenty hundred dollar bills possess a certain quality of convenience and ease in exchangeability which the car does not possess. We might illustrate this again by an analogy from length. If we knew that an automobile was exactly 10 feet long we could, of course, measure the length of a house by taking the automobile and "pacing" the house with it; thus we might find that the house was 5½ automobiles or 55 feet long. But this would be very inconvenient. To measure the length of the house we use a foot rule or, even better, a tape measure. So if we want to buy $2000 worth of cheese we do not take an automobile and say, "Give me $2000 worth of cheese for this." We take twenty hundred dollar bills or, even better, the monetary equivalent of the tape measure, a check for $2000.

The Order of Liquidity

Just as *long* objects can be arranged in an order showing their convenience for use in the measurement of length—with mules, perhaps, at the bottom and tape measures at the top—so *valuable* objects can be arranged in an order showing their convenience for use in exchange, with things like real estate, buildings, furniture, and personal effects at the bottom, ranging up through stocks of standard commodities, ordinary shares, long-term bonds, short-term bonds, call money, and so on through to bank deposits and cash at the top. This is the *order of liquidity.*

The Arbitrary Definition of Money

Things at the top of this scale are called "money"; things at the bottom of the scale, i.e., illiquid things, are what may be bought with money. But just where to draw the line—just what is money and what is not—is difficult to define. Some writers, for instance, include bank deposits on current account in money and some do not. Some writers would confine the term money to mean cash plus bank notes, and some would include even savings deposits. However, strange though it may seem, where the line is drawn does not matter very much. The quality of being money is a matter of degree rather than of kind, and under certain circumstances even things like diamonds or cigarettes behave like money. For purposes of exposition it is often convenient to draw a sharp line at some point in the scale of liquidity and say, "every-

thing more liquid than this is money, and everything less liquid than this is not money." It is important to realize, however, that wherever this line is drawn will be a more or less arbitrary point. Consequently, arguments as to whether a certain thing is or is not money are usually a waste of breath. If we are clear about our definition, whatever it is, we shall not go far astray.

False Definitions of Money

With the above analysis it is at least possible to criticize some attempted definitions of money as not being significant. It is clear, for instance, that in our society gold and silver for most purposes are not money at all, though they were in California around 1850 and still are in a few places. But most people who are likely to read this book would never think of going down to the store with a bag of gold dust, and it would be most doubtful whether they could buy anything with it if they did—the storekeeper would be more likely to call for the police! Similarly, it is not the "redeemability" of paper money which gives it liquidity, for most paper money is not redeemable (i.e., exchangeable into gold or silver at a fixed legal rate); and yet we find no difficulty in buying groceries with it and redeemability in groceries is much more important than redeemability in gold! Nor does the right of legal tender necessarily characterize money—i.e., the legal obligation on the part of creditors to receive the money in settlement of a debt. Most paper money, it is true, is legal tender, but some paper money (e.g., Scottish bank notes) is not, nor are bank deposits; yet the absence of this privilege does not make these assets any less useful as money. Under certain circumstances commodities may become more liquid than "official" money—as were cigarettes in many parts of postwar Europe. We may perhaps therefore venture on a rough definition of money as "those assets which are customarily exchanged for a wide variety of other assets, and which are wanted mainly because of a belief in their continuing ability so to be exchanged."

Money as a Store of Value

The attribute of money as a "store of value" is closely related to its use as a medium of exchange, yet has certain peculiarities. The need for a store of value arises because of a desire to postpone the consumption or use of assets to a future date, usually because of some irregularities in the flow of production or income, or in anticipated needs. Consumption can be postponed, however, only if assets are held in a form which does not depreciate or decay, or at least decays slowly. The property of durability in purchasing power, therefore, has been an im-

portant element in determining the selection of various commodities, or assets as money. The popularity of metals, and especially of the precious metals (those with high value per unit weight), for use as money is closely related with their convenience as a store of value. They are easy to store, do not decay or corrode, and seem to have had a fairly stable purchasing power at least over short periods. No commodity or financial instrument has ever proved completely satisfactory as a store of value, as even those assets which do not decay in a physical sense are almost universally subject to changes in purchasing power. About the closest we could get to a perfect store of value would be a bond or financial obligation with a price-index clause, which would have a constant purchasing power in terms of some price index. Even an instrument of this kind, however, would be no safer than its issuer, and no price index can give a completely satisfactory measure of changes in purchasing power.

Change in Price Level as a Rate of Interest on Money Stocks

The store-of-value aspect of money takes on peculiar importance in times of inflation and deflation when price levels are changing rapidly. When prices are rising, money, along with any other asset which is measured as a constant sum of dollars, declines in purchasing power. Thus, if the price level doubles during a year, $100 will only buy half as much at the end of the year as it would at the beginning. A rise in the price level is therefore equivalent to a *negative* rate of interest, in terms of purchasing power, on all assets whose value is fixed in terms of the monetary unit. Similarly, a fall in the price level raises the purchasing power of these assets, and is equivalent to a positive real rate of interest. If the price level halves during a year, $100 is worth twice as much in terms of its command over commodities at the end of the year than it was at the beginning. Putting the same thing in another way, the owner could spend and consume $50, and still have as much purchasing power at the end of the year as he had at the beginning. It is clear that the expectation of rising prices will lower the willingness to hold money, and that the expectation of a fall in prices will make people more eager to hold money.

THE VELOCITY OF CIRCULATION

If, therefore, we wish to analyze the forces which underly the circulation of money and the volume of exchanges, some method must be devised for describing and defining not only the quantity of money and of other relevant assets but also for measuring the willingness or

"eagerness" to hold money. There are several methods for doing this. A crude but very useful method is through the concept of the *velocity of circulation.* Suppose that we define the money stock of a society, M, in some arbitrary manner. The total volume of *payments* in the society, E, is then defined as the sum of all the quantities of money which change owners in a unit time period. A single payment is a transference of ownership of money from account A to account B. It is recorded as an *expenditure* to A and as a *receipt* to B; the expenditure and the receipt, however, are merely the two "ends" of the same payment. The total of all such payments (which can be reckoned as the total of all expenditures, or as the total of all receipts) is the total volume of payments, E. The velocity of circulation, V, or more accurately the *payments velocity* of circulation, is the ratio of the total volume of payments to the total money stock. That is,

$$V \equiv \frac{E}{M} \qquad \text{or} \qquad E \equiv MV \tag{1}$$

What Does the Velocity of Circulation Measure?

Identities, as we have seen, only have significance if the variables which they contain have a certain independent validity as "parameters of behavior"—that is, as quantities which vary in a regular way with certain aspects of human behavior. The velocity of circulation is such a quantity. It is a rough measure of the willingness to hold money; an increase in the willingness to hold money will diminish the velocity of circulation. It may be thought of as the average number of times in a year that a dollar of the money stock changes hands in a payment. Thus if the "average dollar" changes hands (makes a payment) 20 times a year, and if the total money stock is a million dollars, the total volume of payments will be \$20 million per annum. The reciprocal of the velocity of circulation may be called the "period of turnover"; it is the average period of time which elapses between the successive payments of a single dollar. Thus if the velocity of circulation is 20 times a year, the period of turnover will be 365/20, 18¼ days.

ILLUSTRATION. In order to perceive the truth of this proposition, imagine a society consisting of 30 people in a classroom, each possessing a one dollar bill. The total quantity of money in the society is \$30. Every time a bell rings, each person passes his dollar bill on to another, and receives in turn a dollar bill from another. The acceptance of the dollar bill represents receipts; the giving out of the dollar bill represents expenditure. Suppose the bell rings every minute. The receipts (and the expenditure) of each person are \$1 a minute, or \$60 an hour; the total receipts of the

society are \$30 a minute, or \$1800 an hour, although there are only thirty dollar bills in the room.

Now if the velocity of circulation doubles—that is to say, if the bell rings every 30 seconds instead of every 60 seconds—the period of time for which each individual holds his dollar bill is cut in half. The receipts (and expenditure) of each individual are now \$2 a minute; the receipts of the society are \$60 a minute, or \$3600 an hour. The receipts of the society have doubled, although there is just the same quantity of money in the room, merely because the velocity of circulation has doubled, or, what is the same thing, because people hold on to their money only half as long as they did before!

Suppose again that the bell rings only once a minute, so that the velocity of circulation is again once a minute. Let each individual have two dollar bills instead of one. The receipts of each individual will now be \$2 a minute. The quantity of money has doubled, the velocity of circulation has stayed the same, and consequently the receipts of the society have doubled. Or suppose that the quantity of money doubles—to two dollar bills per head—but that the velocity of circulation is cut in half—to once every two minutes. Then the receipts will be the same as before—\$30 an hour. Evidently, in this very simple case, our formula is correct.

Velocity of Circulation as an Aspect of Human Behavior

The significance of the velocity of circulation concept lies in the fact that it is something more than a convenient statistical ratio; it is a true average of *individual* velocities of circulation, each of which is a figure representing a definite aspect of individual behavior. The average velocity of circulation represents, therefore, a kind of social average of individual patterns of behavior.

In the case of a single individual the ratio $\frac{\text{Money receipts}}{\text{Money stock}}$ may be called his individual velocity of circulation; the ratio $\frac{\text{Money stock}}{\text{Money receipts}}$ in his individual period of turnover. Thus, if an individual with annual receipts of \$10,000 holds a money stock of \$500, his velocity of circulation is 20 times a year. This may be visualized as the number of times a year the reservoir of his money stock would have to be emptied and refilled in making his money payments. His period of turnover is 18¼ days—this may be visualized as the average length of time that a dollar stays in his money stock, or the average interval between receipts and expenditures. The velocity of circulation of the whole system is simply the

weighted average (weighted by the amounts of money held by each individual) of all the individual velocities of circulation.[2]

Effect of a Change in Velocities

The average velocity of circulation is therefore a figure which is determined by individual decisions. Each individual is free to decide how large he wants his money stock to be in proportion to his total receipts; the result of all these decisions is the total volume of receipts. It should be observed that the effect of a change in the individual velocity of circulation is different in the case of an individual from that of society at large. An individual who wants to decrease his velocity or increase his period of turnover does so by increasing the amount of money that he holds, usually—as he regards his receipts as for the moment a given factor—by reducing his expenditures. The effect of his decision on the whole society, however, is not of itself to increase the quantity of money, though that may happen as a secondary effect, but rather to diminish the total of receipts through the diminution of what is the same thing, the total of expenditures, Thus the individual is mainly conscious of his ability to regulate his individual velocity by changing the quantity of money which he holds. The effect of his decision on society, however, will be a change in the total volume of payments, unless there are secondary effects on the quantity of money. Money cannot be destroyed by spending it, for what one spends another gets. This principle will be developed in more detail on pages 81–88.

The "Fisher Identity"

We can now proceed to the consideration of a useful equation or identity known as the "equation of exchange," or, from its author, the "Fisher identity."[3] The Fisher identity is based on the assumption that every payment is payment *for* something and that the dollar value of what is bought and sold with, or for, a payment must be equal to the payment itself.

[2] Let $m_1, m_2, \ldots m_n$ represent the money stocks of the individuals 1, 2, . . . N of a society, and $v_1, v_2, \ldots v_n$ represent their individual velocities of circulation. Let M be the total quantity of money in the society, E the total of receipts (expenditures), and V the average velocity of circulation. Then if $e_1, e_2, \ldots e_n$ represent the individual receipts (expenditures) we have:

$$e_1 = m_1v_1,\ e_2 = m_2v_2, \quad \ldots e_n = m_nv_n, \qquad \text{whence}$$

$$V = \frac{E}{M} = \frac{e_1 + e_2 + \ldots + e_n}{m_1 + m_2 + \ldots + m_n} + \frac{m_1v_1 + m_2v_2 + \ldots + m_nv_n}{m_1 + m_2 + \ldots + m_n}$$

That is, V is the weighted arithmetic mean of $v_1, v_2, \ldots v_n$, weighted by the respective m's. Similarly, it can be shown that the average period of turnover is the weighted *harmonic* mean of the individual periods.

[3] Irving Fisher, *The Purchasing Power of Money*, Macmillan 1911.

The dollar value of anything bought or sold, however, is equal to the price of the thing multiplied by the quantity. Thus, consider a transaction in which 1000 bushels of wheat change hands at \$2 per bushel. The *value* of the wheat is \$2000; the payment which is made for the wheat is two thousand dollars. If, therefore, we represent the price of anything by p, the quantity traded by q, and the amount of money paid by e, we have:

$$pq \equiv e, \text{or } p \equiv \frac{e}{q} \qquad (2)$$

This identity is really the definition of price as the ratio of the money paid for anything to the quantity of the thing bought. If now we define indices for the price level P and the quantity level Q of all things for which payments are made, so defined (as we have seen on page 26) that P represents the total value of all things for which payments are made, then we must have

$$PQ \equiv E \qquad (3)$$

E being the total volume of payments. All that this identity states is that the price of everything exchanged, multiplied by the quantity of everything exchanged, must be equal to the total amount paid for (or received for) everything exchanged. Now putting equations (3) and (1), page 74, together, we have

$$PQ \equiv MV \qquad (4)$$

This is the equation of exchange, or the Fisher identity.

The Fisher identity can also be used in another form known as the income form. If P' represents the price level of output, and Q' represents its volume, $P'Q'$ is the total value of output or the national money *income*. If now we define V', the "income velocity," as the average ratio of individual incomes to the quantity of money, the national income must also be equal to MV'. The income velocity may also be regarded as an expression of the average individual behavior toward money, though it is a concept open to certain objections and must be used carefully. We can, however, write an income identity:

$$P'Q' \equiv MV' \qquad (5)$$

Fisher Identity Models

The Fisher identity is a powerful tool in the analysis of monetary problems—more powerful, probably, than many modern theorists are willing to recognize. It must, however, be used carefully. It separates out the three main elements which affect the general level of prices,

but it does not tell us what determines the magnitude of these elements or how they react on one another. Nevertheless, it forms the basis of a number of interesting models. The simplest model is that which assumes that M is given ($= M_1$) at any particular time by the processes of historical accumulation of the money stock, V is given ($= V_1$) by the psychological propensities of the people, and Q is given ($= Q_1$) by the historical development of productive capacity. This gives us three very simple equations, which when combined with the identity gives an immediate solution for the price level, $P = \frac{M_1V_1}{Q_1}$.

The Quantity Theory

This model, simple as it is, illustrates the conditions under which the "crude" quantity theory of money is true. The quantity theory in its simplest form states that the price level varies in direct proportion to the quantity of money, or $P = KM$. We see immediately from the Fisher identity that this relationship only holds if the ratio $\frac{V}{Q}$ is constant. A rise in the quantity of money will not cause a rise in prices if it is offset either by a decline in V or an increase in Q. The usefulness of the quantity theory as a rough explanation of large changes in the price level lies in the fact that over short periods, at any rate, the possibilities of variation in V and Q are much less than the possibility of variation in M.

An Extended Fisher Model

The Fisher model can be extended by assuming various empirical relationships among the variables. Thus suppose we assume that the quantity of money is itself a function of the price level, and that a rise in prices to a new level will call forth an additional quantity of money, either from the banking system or from the government. Thus we replace the equation $M = M_1$ in the first model by the equation $M = F_m(P)$. Figure 10 shows a graphic solution for this model. The quantity of money, M, is measured on the vertical axis, the price level on the horizontal axis. The curve AB shows the relation between the price level and the quantity of money which each price level calls forth, $M = F_m(P)$. The straight line OB represents the basic identity $M = \frac{Q}{V}$. Q and V are assumed to be given. The slope of this line is $\frac{Q}{V}$; it should be observed that it does not necessarily have a 45° slope, as in the somewhat analogous figure in the basic employment model, Fig. 7, page 54. The point of intersection of AB and OB is the

position of equilibrium. *OR* is the equilibrium price level; *RB* is the equilibrium stock of money. If the price level is below the equilibrium level, say at *OH*, it will call forth a quantity of money *HL*, which will force the price level up to *LG*, and so on. Similarly, if the price level is above the equilibrium it will call forth a stock of money which will force the price level down. The equilibrium, therefore, is stable as long as *AB* cuts *OB* from above; that is, the increase in money stock called forth by a unit increase in the price level must be less than $\frac{Q}{V}$.

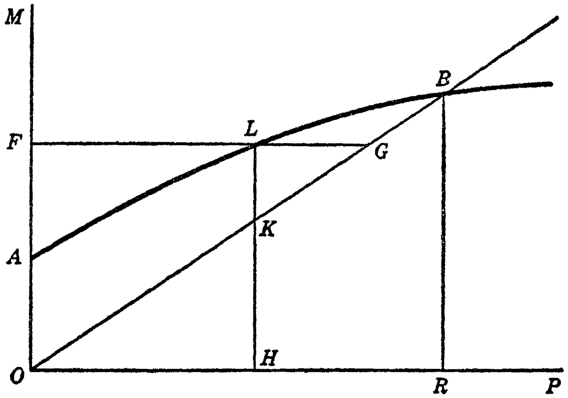


Fig. 10. A Fisher Model

Instability of the Model

This model, simple as it is, throws some light on the mechanics of inflation and hyperinflation. If the money stock is very elastic with respect to the price level, equilibrium will only be established at a high level of prices, if at all. If the elasticity of the money stock with respect to price level is greater than the ratio $\frac{Q}{V}$, no equilibrium is possible, as the lines *AB* and *OB* will not intersect. We are not in a position yet to examine in detail the dynamics of this model. It should be observed, however, that the process of inflation (that is, a positive rate of increase in the price level) is likely to increase the velocity of circulation, as money then loses its ability to be a store of value, and people try to hold as little as possible. Inflation itself, assuming full employment and no increase in *Q*, will increase *V* and lower the ratio $\frac{Q}{V}$. The line

OB will rotate toward the P axis, and the position of equilibrium will move out to higher price levels. The movement toward the equilibrium, therefore, pushes the point of equilibrium away from the present position, like a man chasing a burglar. Similarly, deflation tends to rotate the line OB toward the M axis, and the equilibrium position will fall. It is evident that if equilibrium is to be reached, the rotation of the line OB (that is, the fall or the rise in V) must eventually slow down. Once equilibrium is reached, however, the line OB will swing back again to its original position, and the chase starts all over again! Such a system will obviously be subject to cyclical fluctuations.

Reciprocal Relation of P and Q

An important feature of the economic system to which the Fisher identity calls attention is the reciprocal relation between P and Q. If the total *value* of transactions falls as a result of a change in M or in V, or in both, this fall in value must be reflected *either* in a decline in P *or* in a decline in Q. The more P declines the less Q will decline, and vice versa. This proposition is important in interpreting the behavior of different sectors of the economy in the course of the business cycle. Thus a depression generally results in a fall in the total *value* of the output of all industries. In some industries, however, this fall in the value of output is achieved mainly by a fall in *prices,* in which case output will not be much changed, as in agriculture. In other industries the fall in the value of output is achieved mainly by a fall in *output,* in which case prices are not much changed.

The "Market Identity"

The Fisher identity is essentially a relationship among the "flows" of the system. An identity can also be developed which expresses a price level in a relationship among stocks. This is the "market identity" developed for the case of particular prices in Volume I. It remains to be observed that the identity also holds for any group of commodities or exchangeables, or for all of them. Thus suppose P_a is the price level of all nonmonetary assets, and Q_a is the total physical quantity of nonmonetary assets. These quantities have to be defined, of course, as index numbers, and must be so defined as to make the product P_aQ_a equal to the total value of nonmoney assets, V_a. Then if M is the total stock of money, and r is the preferred liquidity ratio—i.e., the proportion of the total value of assets which people as a whole wish to hold in the form of money—we must have:

$$r \equiv \frac{M}{M + P_aQ_a}$$

that is, $$P_a \equiv \frac{M(1-r)}{Q_a r} \tag{6}$$

This is an extremely useful identity in interpreting movements of price levels, as both M and Q_a can be regarded as historically determined at any one moment and P_a then becomes a simple function of the preferred liquidity ratio or "liquidity preference," r. If r rises, people feel that at existing prices they are holding too little money; they will individually endeavor to increase their money holdings by selling nonmoney assets. A general attempt to sell nonmoney assets, however, will result in a fall in their price level. As the price level falls, the total value of assets falls, and the proportion which the quantity of money bears to this total rises, as the dollar value of the money stock is constant. The fall in prices will go on until people on the whole are satisfied with the higher proportion of money to total assets. Similarly, if r falls, people try to get rid of money by buying other assets. Prices rise, and the total value of assets rises until people are once more satisfied with the amount of money in existence.

THE PAYMENTS MATRIX

In the Fisher identity and the models based on it we consider the payments system as an aggregate. For many purposes, however, it is necessary to break down this aggregate and to consider the individual payment or a smaller aggregate. These can be conveniently expressed in the form of a payments table, or *matrix,* as in Table 4. The various individuals of the society are A, B, C, D. In the table we have considered a society of only four individuals, but the argument can clearly be extended to any number. The individuals are ranged both horizontally and vertically, so that the table has N^2 compartments, where N is the number of individuals. In each compartment we write the total payment *from* the individual in whose row the compartment lies

TABLE 4. THE PAYMENTS MATRIX

	A	B	C	D	Expenditures	Balance of Payments: Hoarding (+) or Dishoarding (−)
A		20	8	15	43	−12
B	5		15	10	30	+ 7
C	16	9		11	36	+ 2
D	10	8	15		33	+ 3
Receipts	31	37	38	36	142	Total 0

to the individual in whose column it lies. Thus in Table 4, 20 represents the payment from A to B, 8 the payment from A to C, and so on. Each payment, therefore, is a receipt to the individual of its "column" and an expenditure to the individual of its row. The total expenditure of each individual is then obtained by adding up the figures in the rows, and the total receipts of each individual is obtained by adding up the figures in the columns. Thus, in the example the total of A's expenditures is 43 and the total of A's receipts is 31.

Total Expenditures = Total Receipts − Money Created

It follows immediately from this figure that the total of expenditures in a closed society is equal to the total of receipts. Whether we add the column of total expenditures or the row of total receipts the result must be the same (142), because these represent simply two different ways of adding up the same list of individual figures. This proposition is only true, however, as long as every receipt and every expenditure is a payment. If money is either created or destroyed by any individual, and if we count the creation of money as a "receipt" and its destruction as an "expenditure," then the above proposition must be modified; its general form is that the excess of all receipts over all expenditures must be equal to the total money created during the period. If money is destroyed, of course, there will be an equal excess of total expenditures over total receipts. Suppose, for instance, in Table 4, that A destroyed 10 units of money. His expenditure would now be 53, but no receipts would be altered; the total of receipts as before would be 142, but the total of expenditures, 152.

From the matrix on Table 4 we can immediately deduce the balance of payments of each of the parties, defined as the excess of receipts over expenditures. Thus A's receipts are 31, his expenditures, 43, so that his balance of payments is −12; B's receipts are 37, his expenditures, 30, and his balance of payments, +7. A positive balance of payments shows by how much the money stock of the party has increased in the period, for receipts represent an addition to and expenditures a subtraction from his money stock. This is called "hoarding" by some economists; thus in Table 4, B has hoarded—that is, increased his money stock by 7 units. Similarly, a negative balance of payments is "dishoarding"—a decrease in the party's stock of money. A has dishoarded 12 units.

It will be observed in Table 4 that the algebraic sum of all the balance-of-payments quantities is zero, or, what is the same thing, the total amount of hoarding is arithmetically equal to the total amount of dishoarding. This necessarily follows if there is no creation or destruction of money, and is a corollary of the proposition that total receipts equal total ex-

penditures. No matter how the receipts and expenditures change, if the total quantity of money in the society is constant, money is merely shifted around from pocket to pocket; and the increase in the pockets of some people must be exactly balanced by the decrease in the pockets of others. This proposition may be generalized in the form that the net hoarding of all individuals in a closed society must be equal to the total amount of money created. If money is destroyed, there will, of course, be net dishoarding. This simply amounts to the proposition that the net change in the sum of the total money stocks of all individuals must be equal to the net change in the stock of money.

Total Hoarding Not Determined by Decisions to Hoard

The proposition is so obvious once stated that it seems almost trivial. Nevertheless, it is astonishing how many people—even bankers—fail to understand it. Thus during World War II there was a large increase in the quantity of money, mainly because the money expenditures of the government were much greater than its money receipts (i.e., there was a deficit in the cash budget). This increase in the quantity of money must, of course, be reflected in the money stocks (balances) of individuals—somebody must hold all the newly created money! The mere fact of the creation of money, therefore, necessitated hoarding —i.e., an excess of money receipts over expenditures—on the part of individuals. Nevertheless, we continually found distinguished members of the financial community making speeches in which they praised the public for the immense amount of "money saved" during the war, apparently oblivious of the fact that these "savings" (i.e., the increase in money holdings) were not the result of financial virtue on the part of individuals, but the result of financial vice on the part of government! No matter what individuals had decided to do about their expenditures and receipts, they could not have helped hoarding during this period while money was being created. This is a truth which we are somewhat unwilling to accept, for we see so clearly that the amount of money that we, as an individual, hoard or dishoard is within limits under our control. It therefore seems absurd, until we look into the matter further, to suppose that the sum of the decisions of individuals in regard to hoarding do not determine the total hoarding of society. Yet such is indeed the case, as can be seen by returning to Table 4.

Decisions to Hoard Decrease Total Payments

We see in Table 4 that A is quite a spendthrift, and that he is dishoarding (i.e., diminishing his money stock) by 12 units in this period Obviously he cannot go on doing this indefinitely, so let us suppose

that he decides to cut down his expenditure in the endeavor to make ends meet. Let us suppose initially that there is no change in the expenditure patterns of the other individuals. We might suppose at first glance that as A has stopped dishoarding and the others have not changed their behavior, that there will be net hoarding in the society. But Table 5 shows what happens.

TABLE 5. EFFECT OF HOARDING

	A	B	C	D	Expenditures	Hoarding (+) or Dishoarding (−)
A		16	4	11	31	0
B	5		15	10	30	3
C	16	9		11	36	−2
D	10	8	15		33	−1
Receipts	31	33	34	32	130	Total 0

Two things should be observed. First, as a result of A's decision to curtail his expenditures, he has, indeed, "balanced his budget," but his decision to do this has resulted in a decline in the total receipts of the society by an amount exactly equal to his contraction in expenditures. This follows inevitably from the proposition that total receipts and expenditures are equal. Second, there is still no net hoarding; hoarding and dishoarding are still equal, the decline in A's dishoarding from 12 to 0 being just counterbalanced by a decline in the hoarding of the others due to the reduction in their receipts which A's decisions involve. Thus we see again that a net decision to increase money stocks (hoard) on the part of individuals does not result in an increase of money stocks unless there is creation of money; it results merely in a decline in money receipts (expenditures). Similarly, a net decision to dishoard (decrease money stocks) on the part of individuals does not result in a decline of money stocks unless there is destruction of money, but results rather in an increase in money receipts (expenditures).

Favorable and Unfavorable Balance of Payments

A positive balance of payments, creating hoarding or an increase in the money stock, is sometimes called a "favorable" balance; and similarly a negative balance of payments, creating dishoarding, is called an "unfavorable" balance. The terminology dates from the days of the mercantilists, and the terms positive and negative are more accurate than favorable and unfavorable. It is not necessarily advantageous to have a positive (favorable) balance of payments, nor is it necessarily disadvantageous

to have a negative (unfavorable) balance of payments in any given period. It is clear that for any individual a positive balance of payments is equal to the amount of his hoarding—i.e., the amount by which his money stocks have increased—and a negative balance is similarly equal to his dishoarding.

And for a Group

From Table 4 an important proposition relating to any group of individuals can be derived—that is, the difference between the receipts of any group from nonmembers and the expenditures of the group toward nonmembers—i.e., the balance of payments of the group—must equal the total net gain in money stocks of the individuals of the group. This is again illustrated in Table 6, derived from Table 4 by simply regarding A and B as a group ("nation"). Their total receipts from "outsiders" (C and D) amount to 43; their total expenditures to outsiders, to 48. The difference (– 5) is the net loss of money from A and B together (A losing 12 but B gaining 7). The reason for this is clear; the payments which are internal to the group (A's to B and B's to A) are common to the group's total expenditures and to the group's total receipts.

TABLE 6. THE BALANCE OF PAYMENTS

	A + B	C	D	Expenditures	Balance of Payments
A + B		23	25	48	–5
C	25		11	36	+2
D	18	15		33	+3
Receipts	43	38	36	117	Total 0

Thus we have

A's total receipts + B's total receipts – A's total expenditures – B's total expenditures = A's hoarding + B's hoarding

Subtracting A's receipts from B and B's receipts from A from the right-hand side, and the identical quantities, B's expenditures to A and A's expenditures to B from the left-hand side of the above equation, we have

(A + B)'s receipts from outsiders = (A + B)'s expenditures to outsiders + the increase in (A + B)'s money stock

The same reasoning can clearly be applied no matter how many individuals are in the group. The important result follows that the individuals of Table 4 can be replaced by groups (e.g., nations) and the

properties of the table remain unchanged, the total of receipts (expenditures) now, however, representing the total of intergroup (international) payments.

Partial Velocities

The question now arises, Can any useful models be set up which will determine the various magnitudes of the payments matrix? The complete determination of the payments matrix, of course, would involve the determination of all the variables of the whole economic system—prices and quantities of goods exchanged as well as payments or values. Nevertheless, it is possible to set up a simple model for the payments system as such by employing the concept of particular or partial velocities of circulation. The partial velocity of circulation from A to B, v_{ab}, is defined as the ratio of A's expenditure to B (or what is the same thing, B's receipts from A) in a given period to A's money stock at the beginning of the period. That is

$$v_{ab} = \frac{\text{A's expenditure toward B in a given period}}{\text{A's money stock at the beginning of the period}}$$

Each partial velocity then is an expression of two aspects of economic behavior. It reflects the attitude of the spender toward money in general, and it reflects his attitude toward the commodity which he is purchasing from the receiver. Thus v_{ab} will probably rise if A decides that he is holding too much money and adopts a more liberal spending policy in general. In this case v_{ac}, v_{ad}, etc., will also probably rise. A rise in v_{ab} may also reflect a shift in A's demand structure toward the things which B is selling and away from the things which C, D, etc., are selling. In this case a rise in v_{ab} may be accompanied by a fall in v_{ac}, v_{ad}, etc. The whole system of partial velocities, therefore, presents a rough picture of the overall structure of demand in a system. It is not, of course, a complete picture, as it takes no explicit account of the relation between purchases and prices. Nevertheless, it gives us a useful halfway house in the development of a model of general economic equilibrium.

Partial Velocities Determine the Payments Matrix

If now the partial velocities corresponding to each expenditure of the system are given, and the total quantity of money in the system is also given, the whole payments matrix can be calculated, on the assumption that all individual balances of payments are zero. Thus suppose, to take the simplest case, that we have a system with only two individuals, A and B. Let a_b be the payment from A to B and b_a the payment from B to A. The payments matrix is shown in Table 7A.

TABLE 7. DETERMINATION OF PAYMENTS MATRIX

	A				*B*	
	A	B	Expenditures		A	B
A		a_b	a_b	A		v_{ab}
B	b_a		b_a	B	v_{ba}	
Receipts	b_a	a_b	$a_b + b_a$			

Suppose now that we are given, as parameters of behavior, the partial velocities v_{ab} and v_{ba}, as shown in Table 7B, and the total quantity of money in the system, M. Suppose this is divided between the two parties, M_a in A's possession and M_b in B's. Then in equilibrium we have

$$a_b = b_a \tag{7}$$

$$a_b = M_a v_{ab} \tag{8a}$$

$$b_a = M_b v_{ba} \tag{8b}$$

whence

$$M_a v_{ab} = M_b v_{ba} \tag{9}$$

we also have

$$M_a + M_b = M \tag{10}$$

Equations (9) and (10) can be solved to give values of M_a and M_b:

$$M_a = \frac{M v_{ba}}{v_{ab} + v_{ba}} \qquad M_b = \frac{M v_{ab}}{v_{ab} + v_{ba}} \tag{11}$$

Then inserting these values in equations (8) we get values for the payments themselves

$$a_b = b_a = \frac{M v_{ab} v_{ba}}{v_{ab} + v_{ba}} \tag{12}$$

Distribution of Money Stock Determined by Relative Partial Velocities

Some interesting features emerge even from this very simple model. The distribution of the money stock of the society between its two parts in equilibrium is seen to be determined by the *relative* partial velocities. If both partial velocities change in the same proportion, there is no change in the distribution of the money stock. If one rises relative to the other, the money stock shifts toward the individual whose partial velocity has risen least.

No matter how many individuals there are in the society, there is normally only one set of payments which will be consistent with any given system of partial velocities, given the quantity of money in the

system and assuming that all balances of payments are zero.[4] The actual solutions become more and more complicated as we increase the number of individuals. Nevertheless certain general principles emerge. An increase in the partial velocities of one set of individuals relative to another set is always likely to push the money stock into the hands of those with the smaller velocities. The money stock must be thought of as a sort of shifting cargo, the distribution of which among its various owners is the result of a complex game of pushball. Individuals, or groups of individuals, will only be successful in diminishing their money stocks if their "push" is relatively stronger than that of others, and will only succeed in increasing their money stocks if their push is relatively weaker than that of others. It can hardly be overemphasized that the distribution of the money stock and of payments is the result of *all* the decisions of *all* the people, and that no one individual can determine the results apart from the decisions of all the others.

The Input-Output Matrix

The payments matrix is affected not only by final demands but also by the demand for intermediate goods as determined by production functions—that is, by how much input of different kinds goes into producing a unit of output. A matrix which shows how much of the output of each sector of the economy becomes input for every other sector is called an input-output matrix. Thus, suppose we divide the economy into four

[4] Suppose we have n individuals, A. B, . . . N. If M_a, M_b, . . . M_n are the equilibrium money stocks of these individuals, the payments matrix is as follows:

	A	B	C	...	N
A		$M_a v_{ab}$	$M_a v_{ac}$		$M_a v_{an}$
B	$M_b v_{ba}$		$M_b v_{bc}$		$M_b v_{bn}$
C	$M_c v_{ca}$	$M_c v_{cb}$			$M_c v_{cn}$
⋮					
N	$M_n v_{na}$	$M_n v_{nb}$	$M_n v_{nc}$		

As for each individual we assume the sum of receipts to be equal to the sum of expenditures, we have n equations of the form

$$M_a v_{ab} + M_a v_{ac} + \cdots + M_a v_{an} = M_b v_{ba} + M_c v_{ca} + \cdots + M_n v_{na}$$

Any one of these, however, can be derived from all the others, so that we have only $n-1$ independent equations. We have, however, a further equation

$$M_a + M_b + \cdots + M_n = M$$

where M is the total stock of money. We have, therefore, n independent equations, sufficient to determine the n unknowns M_a, M_b, . . . M_n. Once the distribution of the money stock is determined, each individual payment can be derived by multiplying the given partial velocity by the appropriate money stock.

sections, one producing durable goods, one nondurable goods, one services and one construction. Here we allocate all the households to one or other of these four sections. Then we could postulate a matrix something like Table 8, showing the input of each sector to every other, including itself, for each sector uses input produced by itself.

TABLE 8. INPUT-OUTPUT MATRIX

	Durables	Nondurables	Services	Construction	Total Input
Durables	30	30	40	20	120
Nondurables	40	40	70	30	180
Services	40	80	100	10	230
Construction	10	30	20	10	70
Total output	120	180	230	70	600

Here the figure in each cell represents the input from the sector of the row to the sector of the column; the figures represent values at constant prices. The sum of the rows is the total input from each sector; the sum of the columns, the output from each sector. In the table we have supposed that input equals output for each sector. If this were not so, then there would be accumulations of goods, or unemployment or overemployment of services. One can suppose, perhaps more simply, that input and output for services must be equal. In this table the figures in each box depend not only on the production functions but also on household demand; to eliminate this feature we could list households as a separate sector of the economy. Each flow of input produces a reverse flow of payments in the payments matrix. Thus, for any sector accumulations or decumulations of product tend to produce decumulations or accumulations of money; as the constantly changing stocks of goods shift from one sector to another, money surges in the opposite direction.

If production functions are assumed to be linear so that a unit increase in input always produces the same increase of output, it is not difficult to calculate coefficients for each cell of the matrix which show how much, for instance, a unit increase in income will increase (or perhaps decrease) the input of the cell in question. This can give valuable insights on how an increase in overall income or output will be distributed among the various sectors of the economy.

QUESTIONS AND EXERCISES

1. After World War II cigarettes came to be used as a kind of money in many parts of Europe. Why do you suppose this happened? Discuss the advantages

and disadvantages of cigarettes in the performance of the various functions of money.

2. Money is sometimes defined as that which is not wanted for its own sake but is wanted only for its power of exchanging for other things. On the other hand, it is also said that money is wanted because it possesses the property of "liquidity," just as water is wanted because it possesses the property of "drinkability." Can these two views be reconciled?
3. What other things besides money can serve as (a) a measure of value, (b) a store of value, (c) a medium of exchange?
4. Suppose the relationship between the stock of money M and the price level P is given by the equation $M = 500 + 10P$. Suppose the velocity of circulation is constant at 20 times a year, and the volume of transactions is constant at 1000 units. What are the equilibrium values of P and M? Solve both algebraically and graphically. Suppose the velocity of circulation rises to 40 times a year. What are now the equilibrium values of P and M? Repeat with $V = 10$ times a year. Illustrate the three solutions on a single graph.
5. Imagine a closed economic system in which the quantity of money was $100,000,000. Suppose that in one year the total volume of payments was $7,300,000,000. Calculate (a) the payments-velocity of circulation, (b) the period of (payments) turnover.

   Suppose that the quantity of money is constant and the payments velocity of circulation falls to 50 times a year. What would be the period of turnover? What would then be the volume of payments? Suppose that the quantity of money rose to $120,000,000 and the period of turnover rose to 10 days. What would be the total volume of payments?
6. The following table represents the expenditures and receipts of a closed society of three individuals, each figure representing a payment from the individual in whose row it stands to the individual in whose column it stands.

	A	*B*	*C*
A		100	150
B	166		58
C	100	156	

   Calculate and tabulate: a. the total expenditure, total receipts, and the net addition to or subtraction from the money holdings for each individual (i.e., hoarding or dishoarding).

   b. the total of all receipts, and show that this is equal to the total of all expenditures.

   c. Suppose that the above table shows the condition in a certain week, week 0. In the next week, week 1, suppose that each individual attempts to "balance his budget" by reducing or expanding his expenditures in order to make them equal to his receipts of week 0. Suppose that the change in expenditures is divided equally between the other two individuals in each case. (Thus, if A diminishes his expenditure by an amount x, take $x/2$ from his expenditure toward B and $x/2$ from his expenditure toward C.) Draw up the payments table for week 1, and repeat the calculations (a) and (b).

It will be found that the individuals have failed to balance their budgets, as their receipts have changed. Continue now to weeks 2, 3, 4, etc., on the same assumptions by which we proceeded from week 0 to week 1. Can you formulate any law governing the movement of the payments, hoardings and dishoardings, and balances? Does the system move toward an equilibrium?

7. a. Repeat Exercise 6c on the assumption that only those individuals who are dishoarding (i.e., have an unfavorable balance of payments) make any adjustments in their expenditure.

   b. Repeat Exercise 6c on the assumption that only those individuals who are *hoarding* make any adjustments in their expenditures. In what way do the results of Exercise 7 differ from the results of Exercise 6c?

8. The following represents another payments table in a society of three individuals:

	A	*B*	*C*
A		64	128
B	64		96
C	128	96	

   Calculate as before the total receipts, expenditures, and balances of payments. Suppose now that C cuts his expenditure in half, paying 64 to A and 48 to B; suppose that he keeps these rates constant. Assume that A and B react to the situation thus created by trying to balance their budgets as in Exercises 6c and 7, making their total expenditures in each week equal to their total receipts of the previous week, and distributing the decline in expenditures equally between the two recipients. Follow the course of the payments table for 5 weeks.

   Make a table showing the total change in money stocks of each individual week by week. Is the system moving toward equilibrium?

9. What will be the equilibrium position of the receipts table of Question 8 after an indefinitely large number of weeks? (Note: Question 9 cannot be answered without knowledge of the summation of infinite geometric series.)

10. If the preferred liquidity ratio is 10 percent and the total quantity of money is $100,000,000, what is the total dollar value of nonmoney assets?

CHAPTER 5

THE FINANCIAL SYSTEM

CLAIMS AND SECURITIES

The analysis of the preceding chapters can be applied fruitfully to the understanding of the financial system, and more especially to the interpretation of the effect of banking on the economic system. What we know as the "financial system" consists of those institutions which are mainly concerned with the creation, destruction, and exchange of "claims." Claims may perhaps be defined as those balance-sheet items which appear on two different balance sheets at the same time, being an "asset" on one and a "liability" on another. Thus, loans, bank deposits, securities of all kinds, annuities, insurance policies, and the like all fall in this category, and the institutions which deal in them—moneylenders, banks of various kinds, discount houses, investment trusts, insurance companies, stockbroking and jobbing firms, stock and bond markets, and so on, are all financial institutions, the greatest of these being, of course, the government.

At the basis of the whole financial system lies a principle which we may call the principle of the creation of claims. Assets can be divided into two great divisions—"things" (including cash) and "claims" (see Volume I, Chapter 15). Every item which features as a claim in the assets of one balance sheet also figures as a liability in the equities of another balance sheet. It follows that when all the balance sheets of a closed society are added together, the total of claim assets is equal to the total of claim liabilities. The total value of the things possessed in a society, and therefore the total value of its individual capitals, is not necessarily connected with the total value of claims. In other words, claims which are assets to one person and liabilities to another may be created without affecting the total value of the capital of a society. When, for instance, a corporation issues bonds, when a building and loan society writes a mortgage, when an in-

dividual makes a deposit with a savings bank, when a commercial bank grants a loan or accepts a deposit, claims are created which figure as an asset in one balance sheet and as a liability in another. The total of all assets is increased, but in the immediate act of the creation of a claim, the total value of things or of capital is not necessarily increased. Of course, as a subsequent result of the creation of claims, the value of things may be increased. If, for instance, a corporation uses the money raised by the sale of bonds to build a factory, the building of the factory will increase the value of the things on the balance sheets and will also result in a net increase in the total of personal capitals (net worths). But the mere creation of claims, in and of itself, does not increase the total of personal capitals.

Why Do We Have Claims?

Why, then, are claims created? Why bother with this complicated structure of debit and credit, which is built up so assiduously and which so often collapses in partial ruin? Obviously because claims (debts and credits) fill a need. In the first place the existence of claims makes it possible to divorce the "real" ownership or "equity" in capital goods from the control of these goods. A man who has a mortgage on his house does not possess in his "net worth" the full value of the house, but within certain limits he is responsible for the house and controls it. A corporation bondholder possesses a certain equity in the general assets of a corporation; but unless the corporation is bankrupted, he has no control over these assets. If there were no claims, each capital good would have to be owned by some individual person, and the great aggregations of goods necessary in the organization of modern industry would be impossible.

The creation of claims is an "expansion event" which expands the total of both assets and liabilities (see Volume I, Chapter 15). Borrowing money on a note, taking out a mortgage, or making a credit purchase are expansion-exchange events for the borrower, which increases his assets but not net worth. Similarly, the payment (destruction) of claims is a contraction event.

Liquid Assets and Commercial Banks

The other need fulfilled by the existence of claims is the need for *liquid* assets, that is, for assets which may be used in the purchase of goods or in the settlement of debt. Generally speaking, there is only one type of financial institution which can create claims of this kind, claims which are not merely assets to their possessors, but liquid assets. This type is the *commercial bank.* When we speak of a "bank" in the following pages, then, we mean a commercial bank, not a savings bank or an investment bank.

Creation and Circulation of Claims

In order to understand the role played by claims in the economic system it is important to distinguish between two distinct but related processes. Just as we have on the one hand the production and consumption of goods and on the other their circulation and exchange, so also we have two distinct operations in regard to claims—one, their creation and destruction, and the other, their circulation and exchange. These processes are often confused in the case of claims because in many instances the production and the exchange of a claim are bound up into a single operation. For the sake of conceptual clarity, however, it is vitally important that these two processes be distinguished.

Creation of Securities

It is not the function of a theoretical work to give an exhaustive list of various types of claims or securities with a description of how they are created and exchanged. Some examples, however, will illustrate the importance of a concept. One of the simplest forms of creation of claims is the issuance of securities, whether bonds or stocks, by a corporation. These securities represent an *obligation to grant* certain benefits on the part of the issuer and a *right to receive* these benefits on the part of the owner. At the moment of their creation, before the securities are sold, we may think of their creation as representing a simultaneous addition to both the assets and the liabilities of the issuing corporation. As a right to receive something, the security is an asset; as an obligation to grant something, it is a liability. The creation of securities always involves a simultaneous creation, therefore, of an asset and an equal liability. The securities do not, of course, usually remain in the possession of the issuing corporation, or there would be very little use in creating them. There is no point in creating securities which are not exchangeable, which is one reason why their creation and exchange are so often confused. Any individual could write himself out a promissory note for $1 million, and so inflate his assets and liabilities by that amount, but the action would not be very significant unless he could find a buyer for the note! Similarly, corporate securities are issued in the hope that they will be exchangeable for other assets, and ultimately for goods.

Exchange of Securities

Once securities are "born," then, they almost always enter some kind of a market. In the case of the stocks or bonds of recognized companies, this market is highly organized, competitive, and speculative. A security once

launched may change hands many times before it disappears. The prices of securities are determined, much as any other prices (see Volume I, Chapter 7), by the relative quantities of the different kinds and by the relative preferences for them. If the preference for any particular kind of asset declines relative to others, that is to say, if people become relatively less willing to hold it, its price will fall and its value decline relative to others until people can be found who are willing to hold the quantity that exists in the market. Similarly, if the quantity of any particular kind of asset rises, in the absence of any change in preferences, the price will have to fall until enough people can be found who are willing to hold this increased quantity.

THE BANKING AND FINANCIAL SYSTEM

Savings Banks

The banking system is so important as an instrument for the creation of claims, and especially of liquid assets, that it is necessary to examine its operations in some detail. We may dismiss fairly briefly those elements of the banking system outside the commercial banks, as not presenting any difficult theoretical problems. Savings banks create claims in the form of savings deposits, which are liabilities to the bank and assets to the depositor. When a savings bank "sells" $100 of these deposits to the depositors for cash (or, say, for commercial bank deposits), in the first instance the savings bank increases its assets by $100 in cash or commercial bank deposits and increases its liabilities by $100 in its own deposits. The savings banks, however, will normally exchange most of this cash for earning assets of some kind, such as bonds. If no new bonds are being issued, the operations of savings banks may force up the price of bonds, thus, as we will see more clearly later, lowering the rate of interest. If the creation of new bonds is elastic with respect to the interest rate, the operations may increase the quantity of bonds rather than affect their prices.

Investment Trusts

Investment trusts and holding companies are a significant element in the financial system, but again present few problems of a general theoretical nature. The device of incorporation permits the "pyramiding" of claims through the development of intermediary corporations whose assets consist of the claims of other corporations and which themselves issue securities. Thus an investment trust may hold as assets the securities of a great variety of other corporations, and will itself issue shares which

it can persuade individuals, or other corporations, to hold. By this means the risks of the individual holder are spread, and people may be persuaded to hold a greater volume of securities than otherwise would be the case. Abuse of the holding company device, especially in the 1920s, led to fairly severe legal restrictions on its use.

Insurance Companies

Insurance companies, especially life insurance companies, have become an increasingly important element in the financial system. Here we must distinguish between two forms of insurance. On the one hand, "term" or pure service insurance consists essentially of the purchase of "protection" against certain possible losses or calamities by making a "bet" with the insurance company to be paid off if the feared calamity occurs. In this type of insurance no equity is created for the insuree, except in so far as premiums are paid in advance and have unexpired value; the insurance transaction does not involve *saving*. Most life insurance policies, however, are not of this kind, but involve the payment of larger premiums than would be necessary to buy pure protection, and so build up the net worth of the insured persons. Insurance companies thereby become important financial intermediaries, holding as assets large amounts of securities of other institutions and showing as liabilities the net worth or equities accrued to their policyholders.

Commercial Banks

None of these institutions, however, has the impact on the economic system of the commercial banking system, and to this we must now turn.

Our first task is to clear away any possible illusions concerning the nature of a commercial bank. A bank is not, primarily, an institution for keeping money locked up in a safe place. It is true that the first bankers were goldsmiths, to whose care people entrusted their surplus stocks of gold, and it is true that to this day most banks have safety deposit boxes in which valuables can be deposited for safekeeping. But this is only a minor part of a bank's business, and a "bank deposit" does not mean a sum of money locked up in a bank's strong room. Another illusion concerning banks is that a bank is an institution which "lends your money to other people." It is true that the primary function of a bank is to grant and to receive loans, and that when one person lends money to a bank, the bank is thereby enabled to make loans to others. But a bank is much more than a loan broker, and it does more than act as an intermediary between those who wish to lend and those who wish to borrow.

Deposits

The loans which are made to a bank are called "deposits." The name is an unfortunate one, for it calls up a picture of a little pile of money resting in the bank's vault. But a deposit is not money given to the banker for safekeeping; it is a *loan* to the banker, given to him under certain peculiar conditions. Chief among these conditions is that the loan shall be repayable *on demand,* i.e., at any time the depositor wishes to have it repaid. A bank deposit, that is to say, is a peculiar kind of "security." It represents a promise made by the banker to pay the owner of the deposit any sum, in legal money, up to the amount of the deposit, at any time when the depositor may walk into the bank and demand it. There are two principal forms of bank deposit. Demand deposits, or current deposits, represent an obligation of the bank to pay the sum in question immediately on the demand of the owner. Time deposits, or savings deposits, represent an obligation of the bank to pay the sum in question a certain time (usually three days) after notice has been given. We shall not be concerned with the specific consequences of this distinction, as these are matters of detail rather than of principle.

Purchase and Sale of Deposits for Cash

One of the things a bank does, then, is to receive cash in return for deposits and to pay out cash in settlement of deposits. If I take $100 to the bank in cash (legal tender) and pay it into my deposit, this is in a sense a "purchase" of a bank deposit worth $100. The result of the transaction is that the bank now has $100 in cash which it did not have before, and it also has an obligation to pay me $100 which it did not have before. On the other hand, I have $100 less cash than I had before, but I possess a promise by the bank to pay me $100. In effect, I now possess a kind of "security" worth $100, in the form of a bank deposit. Similarly, when I withdraw $100 of my deposit in cash, the bank as a result has $100 *less* cash than it had before, and I have $100 *more* cash. But the bank no longer owes me $100, as I have given up my deposit.

Bank Notes: Similar to Deposits

Another thing a bank may do is to issue *notes.* This privilege is by now largely of historic interest, as far as the ordinary banks of England and the United States are concerned, for the privilege of note issue has been gradually confined to the Bank of England in the one case and the Federal Reserve banks in the other. In the nineteenth century, however, nearly all banks issued notes, and even now in some countries—e.g.,

Scotland and France—note issue is still an important part of a bank's business. A bank note is an obligation much like a deposit. A deposit, however, is an obligation of the bank to a particular person, whereas a note is an obligation of the bank to anyone who possesses the note. That is to say, a bank deposit of $100 is a promise, made by the bank, to pay a specific person—Mr. Jones—$100 on request. A bank note of $100 is likewise a promise on the part of the bank to pay $100—to the person who owns the note and presents it to the bank. When a bank pays out legal tender money for a bank note, the process is called "redemption," for when that is done the bank "redeems" its promise.

Other Methods of Acquiring Deposits

An individual may acquire a deposit in ways other than purchase with cash. Indeed, the commonest way for an individual to acquire a deposit is by paying in a check. Another common way is by selling something to a bank, such as a bond or a promissory note. Indeed, the key to understanding the banking process lies in the analysis of what happens when a bank buys something, whether it is a bond, a mortgage, or a promissory note. It should be observed that "making a loan" on the part of a bank is essentially similar to "buying a security." What happens when a bank makes a loan is that it "buys" from the borrower a promissory note, promising to pay the bank a certain sum on a certain date. The only difference between a bond and a promissory note is that a bond generally represents a promise to pay a series of sums at a series of dates, whereas a promissory note generally, though not always, consists of a promise to pay a single sum on a single date.

How the Purchase of Assets Creates Deposits

Now, what happens when a bank buys anything—say, a bond from Mr. Smith? The ownership of the bond is transferred from Mr. Smith to the bank and in the first instance the bank gives Mr. Smith a *deposit* equal to the value of the bond which has been purchased. That is, Mr. Smith gives the bank a bond and receives in return a promise to pay—a deposit—equal to the value of the bond. This may seem a little strange until we recall that a bond is also a promise, so that the exchange of one promise for another is not unreasonable. Notice, however, what has happened in this transaction. The number and value of assets has increased, for a promise to pay is an asset whose value can be measured in dollars. It is almost as if Mr. Smith said to Mr. Jones, "You lend me a million dollars and I'll lend you a million dollars." In that event Mr. Jones and Mr. Smith might each give to the other an I.O.U. for $1 million. Each

would have a piece of paper worth, nominally, $1 million. Two million dollars' worth of securities would have been created, apparently by the mere stroke of a pen. Something of this kind happens when a bank grants a loan. By that act securities (promises to pay) are immediately created to the value of twice the amount of the loan. Suppose, for instance, that a bank makes a loan of $5000 to Mr. Robinson for six months. The moment after the loan has been made, the bank owns a security worth $5000—Mr. Robinson's promise to pay the bank $5000, or rather more if we take interest into account, in six months. Mr. Robinson now owns a security also worth $5000—a deposit at the bank, or a promise on the part of the bank to pay Mr. Robinson $5000 on demand.

Why Deposits Are Regarded as Liquid Assets by Their Owners

However, there seems to be a difference between this case and the case of the two friends who loaned each other $1 million. The transactions of Mr. Jones and Mr. Smith are purely fictitious, for neither could fulfill his promise alone. They could cancel out their obligations, but one could not, presumably, pay off his obligation without the other at the same time paying off his. In the case of the banker this is not so. The banker can fulfill his promise to pay Mr. Robinson $5000 long before Mr. Robinson fulfills his promise to pay $5000 to the banker. If Mr. Robinson wishes, the banker will carry out his promise immediately the loan has been made, and give Mr. Robinson $5000 in cash. This will extinguish Mr. Robinson's deposit, for the result of this transaction is that $5000 in cash is transferred from the banker to Mr. Robinson, and a $5000 "security"—the bank deposit—has been destroyed. Suppose Mr. Robinson takes his $5000 *in cash* as soon as he has made his promise to pay this sum back to the banker, with interest, in six months. Then the net result of these transactions is that $5000 in cash has changed owners, and a security of $5000 (Mr. Robinson's debt to the bank) has been created. This is what happens when anyone makes a loan. I do not have to be a banker to lend $5000 in cash to Mr. Robinson, and he might as well have obtained it from me, a private individual, as from a banker—if he could have done so. By a private loan of cash a security is also created—the promise to pay back the cash in a given time. The thing to be explained, which is peculiar to bankers, is how a banker, in making a loan, contrives to create not one security but two—one, the loan itself, and the other, a bank deposit which is not immediately redeemed. For normally Mr. Robinson will not transform his $5000 deposit into cash immediately. He may transform some of it into cash, but most of it he will probably transfer to other people by means of the instrument known as a *check*.

The Transferability of Claims

How can deposits be transferred from one person to another? The transfer of deposits is a special case of an important general principle, which we may call the principle of the transferability of claims. If A owes a debt to B, he can extinguish it by passing on to B a claim which A has against a third party, C, provided B is willing to accept this claim.

TABLE 9. THE TRANSFERENCE OF CLAIMS

A		B		C	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1. From C $100	to B $100	From A $100			to A $100
2.		From C $100			to B $100

Thus suppose A owes B $100 and C owes A $100. This situation is reflected in the balance sheets in item 1 of Table 9. Now if A can persuade B that a debt from C is just as good as one from A, he can pay his debts (i.e., cancel out his assets against his liabilities) by giving C's note to B. The situation then is shown in line 2. It may be observed that if in the above situation C had owed $100 to B, all the debts could have been cancelled out against each other. This is called "clearing."

Bills of Exchange

A bill of exchange is an order written by one person, A, ordering another person, B, to pay a given sum to a third person, or to the holder of the document, C. Suppose that B owes a debt to A—i.e., that A has a claim on B for, let us say, $100—and A owes $100 to C. A then writes what is in effect a "letter" to B, or, as we say, he "draws" a "bill of exchange" on B, ordering B to pay C a sum of $100, usually in two or three months' time. A then sends this document to C. Its possession gives C a claim on B for the sum stated. If, therefore, C has a debt to pay to another person, D, he can pay his debt by "endorsing" the bill and passing it on to D. The endorsement consists now of C's signature on the back of the bill. The signature, however, really symbolizes another letter to B, saying in effect: "Dear B, please pay to D the $100 which you are ordered to pay to me. Yours sincerely, C." D now owns a claim on B for $100, which he can pass on to E if he wishes, and so on, until finally the time comes when the initial order has to be fulfilled. B then pays $100 to the person who owns the bill, and the bill disappears from circulation. This method of settling debts is but little used now,

except in international trade. At certain stages of economic development, however, it is common in internal transactions. Everywhere, in fact, it tends to be a stage in the development of banking.

The Checking System

The commonest method now of settling debts in the more advanced systems is by *check*. The inconvenience of a bill of exchange is that it gives the owner a claim for the sum stated *only at a specific date* in the future. Obviously one would normally be more willing to accept in settlement of a debt a document which entitled the owner to receive the sum stated *at any time*, i.e., on demand. But this is precisely the nature of a check. A check, in fact, is a "bill of exchange" drawn on a banker. A bank deposit is a debt, payable by the banker on the demand of the depositor, who is the creditor. If A owns a bank deposit, he owns a promise, made by the banker, to pay him a sum on demand equal to that deposit. If, therefore, A owes C a sum of $100, and has a deposit with Banker B of an amount not less than $100, he can send a letter to the banker which might perhaps read: "Dear Banker, Please pay Mr. C the sum of $100 whenever he wants it. Yours sincerely, A." This letter is a check. In practice it has been formalized a little, but if a check is examined it will be seen to follow closely the form of a letter to the banker. It would be possible for C to endorse this check and pass it on as it stands to D, and for the check to circulate just as bills of exchange have occasionally circulated. This may happen sometimes, but it is rare. Most people, on the receipt of a check, take it to a bank. They may compel the banker to fulfill immediately the obligation which the check implies, that is, to exchange the check for cash. Or they may prefer to exchange the check for a bank deposit of the same value.

Bank Deposits as Money

This fact—that a bank deposit can be transferred easily from one person to another by means of a check—explains why bank deposits exist at all. A bank deposit is a promise, made by the banker, to pay cash to the owner of the deposit any time he wants. Why, then, does not everyone call on the bankers to fulfill this promise? Why does not everyone substitute cash for the sum of his bank deposit? In answering these questions we must first ask, "Why do people want cash?" The answer is, "To pay for things," or, what is practically the same thing, "To settle their bills," i.e., to fulfill their own promises to pay. But, as we have seen, Mr. Jones can fulfill a promise to pay, say, $100 to Mr. Robinson by writing him a check. Indeed, this may be a more convenient way of paying Mr. Robinson than paying him in cash, for cash may be lost or stolen on the

way, while the check is no use to anyone but the person for whom it is intended. For making many payments, therefore, checks are a more convenient form of payment than cash. This is not true of all payments; generally speaking, small irregular payments to persons with whom we are not very well acquainted are best made in cash, and large regular payments to persons with whom we are acquainted are most conveniently made by check. Most people who have payments of this latter kind to make will wish to have some proportion of their resources in the form of bank deposits and some proportion in the form of cash. Bank deposits, that is to say, are a liquid form of property, like cash. Indeed, for certain purposes they are actually more liquid than cash, and therefore more entitled to be called "money." We should not be surprised at this if we remember that all money, even cash, is essentially a promise to pay goods and services up to the appropriate value. If I own \$100 in cash the money signifies that I have a right to receive, on demand, \$100 worth of goods and services. If I own a \$100 bank deposit, the deposit signifies that I have a right to receive, on demand, \$100 in cash—which is the same thing as a right to receive \$100 in goods and services. It is little wonder, then, that cash and bank deposits are such close substitutes for each other.

THE THEORY OF BANKING BEHAVIOR

The Balance Sheet of a Single Bank; Events Table

Perhaps the simplest way to illustrate these principles is by arithmetical examples. Imagine a system in which there is only a single bank—let us call it The Bank. Then in Table 10 we see a highly simplified events table (see Volume I, Chapter 15) showing the possible effects of the introduction of a bank into a system which previously did not possess one. We assume a highly simplified balance sheet, with assets of The Bank consisting only of loans and investments (securities held) S, and reserves R, (which in this case we suppose to consist of legal tender money) and with liabilities of The Bank consisting only of bank deposits, D, and The Bank's capital or net worth, C. The balance-sheet identity then gives us

$$S + R = D + C \qquad (1)$$

In this system we suppose a constant stock of legal tender money N, of which M is held outside The Bank and R inside, so that we have another identity:

$$R + M = N \qquad (2)$$

In Table 10 N is assumed to be 2000. Then we have two behavioral equilibrium conditions:

$$r = \frac{R}{D} = \bar{r} \tag{3}$$

$$h = \frac{M}{D} = \bar{h} \tag{4}$$

where $\bar{r}$ and $\bar{h}$ are constants. r is the reserve ratio of The Bank; the behavioral rule is that if r is less than $\bar{r}$, The Bank will contract S by selling securities or calling in loans, and if r is greater than $\bar{r}$, The Bank will expand S by making new loans or buying securities. h may be called the cash preference of the public. The behavioral rule is that if h is less than $\bar{h}$, people will withdraw cash from The Bank (that is, buy cash with deposits), and if h is greater than $\bar{h}$, people will take cash to the bank and buy deposits.

TABLE 10. EVENTS TABLE OF SYSTEM WITH A SINGLE BANK

	Period						
	1	2	3	4	5	6	7
Bank assets							
Loans and investments (S)				6000	6000	5500	6000
Reserves (R)		1000	1500	1500	500	500	500
Total		1000	1500	7500	6500	6000	6500
Bank liabilities							
Deposits (D)			500	6500	5500	5000	5000
Capital (C)		1000	1000	1000	1000	1000	1500
Cash outside bank (M)	2000	1000	500	500	1500	1500	1500
$r = R/D$ ($\bar{r} = .1$)			3.0	.23	.09	.1	.1
$h = M/D$ ($\bar{h} = .3$)			1.0	.075	.27	.3	.3

Period 1 is before the establishment of The Bank, with 2000 in cash in the hands of the public. Between 1 and 2 The Bank is founded by people subscribing 1000, which becomes the first reserve. Between 2 and 3 we suppose that people move towards a preferred cash ratio by taking 500 to The Bank and buying deposits, which adds 500 to both reserves and deposits. The Bank finds itself with a high reserve ratio (excess reserves), so makes 6000 in loans and investment, adding 6000 also to its deposits, which brings us to period 4. This, however, has lowered the cash ratio h to .075, so the public withdraws cash from The Bank (1000), lowering reserves and deposits by 1000, which brings us to period 5. The Bank now finds that its reserve ratio is too low, so it contracts S by 500, by selling securities or calling in loans, which diminishes both S and D by 500 and

brings us to period 6. As we have assumed $\bar{r} = .1$ and $\bar{h} = .3$, the system is now in equilibrium. Period 7 is added to show how a bank earns profits; we suppose interest accrues on loans and investments equal to 500; this raises S by 500 and also raises C by 500, leaving r and h unchanged. If now The Bank pays a 500 dividend to its owners, we move to the position of period 5, and if then the accrued interest is paid to The Bank by extinguishing deposits (writing a check to the bank itself), we move to the position of period 6 again.

From equations (1) to (4) we can calculate the equilibrium position. From (2) (3) and (4) we have

$$D = \frac{N}{\bar{h} + \bar{r}} \tag{5}$$

$$R = \frac{\bar{r}N}{\bar{h} + \bar{r}} \tag{6}$$

$$M = \frac{\bar{h}N}{\bar{h} + \bar{r}} \tag{7}$$

and now from (1)

$$S = \frac{(1 - \bar{r})N}{\bar{h} + \bar{r}} + C \tag{8}$$

We see immediately that the smaller the reserve ratio, $\bar{r}$, and the smaller the preferred cash ratio, $\bar{h}$, the larger will be the amount of deposits, D, created and the larger the amount of bank loans and investments, S.

Graphic Solution

A graphic solution to this model is shown in Fig. 11. Here we measure M to the right, R to the left, D upwards, and C downwards from the origin, O. The line OM_e shows the relation $M = \bar{h}D$

$$\text{Thus } h = \frac{D_eM_e}{OD_e} = \tan D_eOM_e \tag{9}$$

The line OR_e shows the relation $R = \bar{r}D$

$$\text{Thus } r = \frac{R_eD_e}{OD_e} = \tan R_eOD_e \tag{10}$$

We now move the horizontal intercept R_eM_e upwards until $R_eM_e = N$, the total quantity of cash. At this point OD_e is the equilibrium quantity of deposits, D, R_eD_e the equilibrium quantity of reserves, R, D_eM_e of cash outside the bank, M. Then, if we draw R_eS at a 45 degree angle from

R_e to cut OD_e in S_e, and if $OC_e = C$, then the total assets of The Bank are C_eD_e $(= C + D)$, and the total liabilities, C_eS_e $(S) + S_eD_e$ (R).

To construct the point D_e, draw OA so that $R_eOA = D_eOM_e$, make $OR = N$, and draw the perpendiculars RA to OA, and AD_e to OD_e.

A System with More Than One Bank

Now if instead of one bank our system has a number of banks, what will be the effect of an expansion of loans by a single bank—let us call it the Podunk Bank? The effect will be similar in kind to the case of the

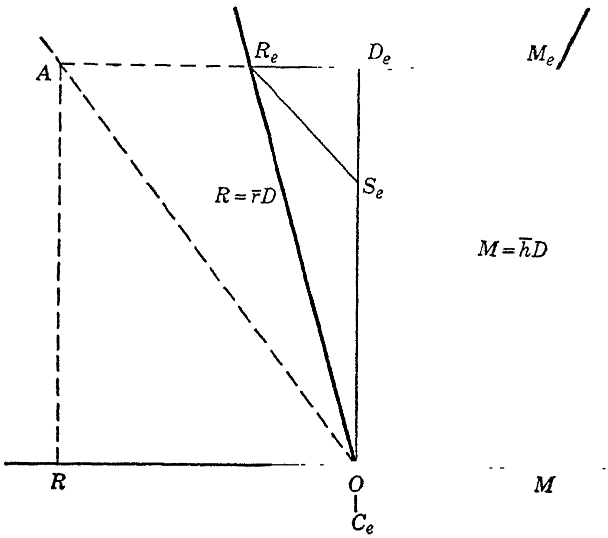


Fig. 11. Equilibrium of a Bank

one-bank system which has just been discussed. If the Podunk Bank increases its loans by \$500, it may expect to have some increase in its deposits, coupled with some loss of its cash. The loss of cash, however, can take place for a reason additional to the one already given. The Podunk Bank will lose cash not only because some of the people, whose deposits have been increased as a result of the loan, will want to exchange part of their increased deposits for cash but also because some of them will transfer these deposits, by check, to people who have deposits *with other banks.*

EFFECT OF A CHECK ON ANOTHER BANK. What happens, then, when Mr. Smith, who banks with the Podunk Bank, gives Mr. Robinson, who banks with the Toonerville Trust, a check for \$100? This check is a "letter" to the Podunk Bank requesting it to pay Mr. Robinson \$100 cash on

demand. Suppose that Mr. Robinson does not want $100 cash, but wants to add $100 to his deposit with the Toonerville Trust. He gives this check to the Toonerville Trust, with his signature upon it, i.e., his endorsement. His endorsement, as we have seen, is an abbreviated form of another letter, also to the Podunk Bank, saying, "Will you please pay this $100 to the holder of the check. Yours sincerely, J. Robinson." That is to say, as he has given it to the Toonerville Trust, the check with his endorsement is now an obligation of the Podunk Bank to pay $100 to the Toonerville Trust. Now, the Toonerville Trust will not be content to hold this promise in its hand; it will go to the Podunk Bank and collect the $100 in cash, and the deposit will finally disappear.

Loss of Cash to Other Banks Prevents Unilateral Expansion. If, then, these two banks are the only two banks in the system, and if the Podunk Bank increases its loans to Mr. Smith by $500, the results will be something like the following: The Podunk Bank will gain some deposits—say, $200—in so far as Mr. Smith retains his new deposit or pays it out in checks to people who bank with the Podunk Bank and who also retain their new deposits. The Toonerville Trust will also gain deposits—say, by $250—for Mr. Smith will pay some of his $500 to people who bank in Toonerville, and who will also wish to retain some of their deposits. The Podunk Bank will lose a certain amount of cash to its depositors—say, $50—and the Toonerville Trust will likewise lose some cash to depositors, assuming constant preferred cash ratios, for the total value of deposits has now increased and people will withdraw some cash to balance this. The Podunk Bank will also pay out cash to the extent of $250 to the Toonerville Trust, corresponding to the increase in the Toonerville Trust's deposits, for that has been due to the paying in of checks on the Podunk Bank which the Toonerville Trust will collect. By making a loan of $500, then, the Podunk Bank will gain $200 in deposits and lose $300 in cash. It is evident that in this case the ability of the Podunk Bank to increase its loans *all by itself* is very limited; much more limited than in the case of The Bank of our one-bank system, for there a loan of $500 might only result in a $50 loss of cash.

But All Banks May Expand Loans Together. Now, what happens in our two-bank system when *both* banks increase their loans? In the absence of any loans by the Podunk Bank, a $500 loan by the Toonerville Trust may be expected to increase its deposits by $200, decrease its cash by $300, and increase the deposits of the Podunk Bank by $250. Suppose, however, that *both* the Podunk Bank *and* the Toonerville Trust increase their loans by $500. The Podunk Bank will gain $200 in deposits from its own loan, plus $250 in deposits from the Toonerville Trust loan, or $450 in all, while it will lose $50 cash to the public. Likewise, the Toonerville Trust will gain $450 in deposits and lose $50 to the public. Neither

bank will pay cash to the other, for each will have a claim of $250 on the other, which can be "cleared." In this example we have assumed that the banks are the same size and are symmetrical in their relations. If the Podunk Bank is a large one and the Toonerville Trust a small one, then a given loan by the Toonerville Trust would bring about a large transfer of cash to the Podunk Bank, whereas the same loan by the Podunk Bank would bring about only a small transfer of cash to the Toonerville Trust.

It is clear, then, that while the lending ability of a single bank, in a two-bank or a many-bank system, is very limited, the ability of the system as a whole to expand its loans is considerable, provided that all the banks do it together. A many-bank system is rather like a number of balloons tied together with string; if one balloon tries to get away from the others the strings will bring it back, but all the balloons can rise together without difficulty. The "string" in the case of a banking system is the loss of cash reserves. One bank expanding loans disproportionately will lose reserves to the other banks, but if all banks expand together they will all lose reserves to each other, which means of course that no bank loses reserves on balance except to the public.

EFFECT OF DECREASE IN LOANS. An exactly similar process, in reverse, happens when a bank decreases its loans. If the Podunk Bank decreases its loans by $500 it will lose some deposits, and the other banks will also lose some deposits; but it will gain cash both from the public and from the other banks. Its reserve ratio will therefore rise. If its reserve ratio is persistently above the limit which it wishes to set, it will try to expand the volume of its investments by granting loans or buying securities. If it finds its reserve ratio falling below the limit, it will contract the volume of its investments by not renewing loans as they fall due, or by selling securities, and so will increase its holdings of cash.

The same principles can be applied to a many-bank system, as shown in the appendix to this chapter. The general principle emerges that the smaller the reserve ratios and the smaller the proportion of the total money stock people wish to hold in the form of cash, the greater will be the equilibrium volume of deposits for a given total cash stock.

CENTRAL BANKING

The Central Bank

In the foregoing argument we have assumed that the individual banks of a system hold their reserves in the form of cash. In a developed system this is not usually the case, for the individual banks will hold their reserves in the form of deposits at a *central bank*. A central bank, such as a Federal Reserve bank, the Bank of England, the Bank of France, or the Bank Deutscher Länder, has several tasks to perform. It is a bankers'

bank; it makes promises to pay money to bankers (deposits) and also receives from bankers promises to pay money to it. It is the government's bank; it holds the government's deposit, and may make the government loans. It is usually entrusted by the government with a certain responsibility for the management of the monetary and banking system of the country. It frequently holds the national stock of gold, and usually issues notes which form part of the stock of legal tender.

Member Bank Reserves Are Deposits at the Central Bank

When individual banks find that they have claims on each other, which in the absence of a central bank would have to be paid in cash, they can pay them by transferring deposits at the central bank. If, for instance, the Toonerville Trust found that at the end of a week it possessed checks on the Podunk Bank for $5000, and the Podunk Bank possessed checks on the Toonerville Trust for $6000, these claims could be settled by a transfer from the Toonerville Trust to the Podunk Bank of a $1000 deposit at the Federal Reserve bank.[1]

A Central Bank Can Create Bank Reserves

The significance of a central bank lies mainly in the fact that it can create *bank reserves,* much as an individual bank can create deposits, by the process of making loans to the individual member banks, or by the purchase of other securities.

REDISCOUNTING. It can make loans by the process known as "rediscounting." A banker will go to his Federal Reserve bank with "commercial paper" of a special kind—that is to say, with a document which gives the holder a right to receive a certain sum on a certain date from some firm or individual. He then hands over this paper and receives in return an equivalent deposit with the Federal Reserve bank. This commercial paper is supposed to be the result of some genuine commercial transaction. Mr. Smith, for instance, may buy wheat for shipment to Europe and know that he will be paid for the wheat within 60 days. He has to pay for the wheat he has bought, we will suppose, on the day when he buys it, and may not have the cash in hand. So he borrows from a bank for 60 days. That is, he gives the banker a document stating that he, Mr. Smith, promises to pay the banker, say, $5000 in 60 days from today. In return the banker gives him a deposit for a little less than $5000, say, $4950. If now the banker is running short of reserves he may take this document to

[1] In the English banking system the Bank of England is the only bankers' bank, and virtually all transfers of reserves among the member banks are carried out by transferring the ownership of deposits at the Bank of England. In America the Federal Reserve System has not yet achieved this position, for a number of private banks, in New York City especially, hold deposits which belong to the smaller banks, and so act as bankers' banks.

the Federal Reserve bank and rediscount it—give it to the Federal Reserve bank in return for a deposit which is just as good as a cash reserve.

ADVANCES. A Federal Reserve bank can also increase a bank's reserves by making "advances" to it; i.e., by giving it a deposit in exchange for a promise on the part of a member bank to pay the Federal Reserve bank an equivalent sum on a given future date.

OPEN MARKET OPERATIONS. Commercial bank reserves are also changed, as explained below, by purchases and sales of securities by a central bank in the open market.

How the Central Bank Controls Credit

THE REDISCOUNT RATE. It is evident, then, that the central bank can have a great deal of influence on the conduct of the member banks (and therefore on the whole volume of loans and deposits), and through these, on the whole structure of incomes, prices, and interest rates. It exerts this control in a number of ways. First, it may change the rate of interest at which it is prepared to lend. This is the "bank rate" of the Bank of England, the rediscount rate of the Federal Reserve banks. In the case of the Bank of England this power was very important, for the bank rate was an effective weapon in changing the rates of interest at which the *member* banks were prepared to lend. This was partly because the Bank of England at one time acted not only as a central bank but also as a private bank, i.e., it competed to some extent in the market with the other banks in granting loans to private firms and individuals. Consequently, the rate at which the Bank of England lent "led the market," and the custom grew among the other banks of lending at this rate, less a fixed amount. This habit persisted even after the Bank of England ceased to compete with the private banks. When the Bank of England raised the bank rate, the other English banks also raised their rates of interest, and this helped to choke off loans. Similarly, when the Bank of England lowered the bank rate this helped to raise the volume of loans.

In the Federal Reserve System the rediscount rate is not such an important instrument of policy. It does have an effect, however, in encouraging or restraining borrowing by the member banks. If, now, the central bank can raise or lower the rate of interest at which people can borrow from the banks, it will clearly be able to affect the volume of bank deposits. If it raises the rate of interest, borrowing will be discouraged. Loans that fall due will not be renewed, loans that might have been made will not be made, and the volume of both loans and deposits will shrink. Similarly, if it lowers the rate of interest, borrowing will be encouraged, and both loans and deposits will increase.

OPEN MARKET POLICY. Another weapon of the central bank, of great

importance nowadays, is what is known as "open market policy." This means the direct manipulation of the securities market by buying and selling securities on the open market, in order to raise or lower the price of securities and in order to raise or lower the volume of central bank deposits (i.e., bank reserves). Thus when a central bank buys securities, it pays for them by creating central bank deposits; that is, it writes up in its balance sheet an asset (the security bought) and a liability (a deposit owned in the first instance by the previous owner of the security). If the seller was a bank the total of bank reserves is immediately increased by the amount of the reserve bank deposit. If the seller was not a bank (say a private individual) he is likely to sell his central bank deposit to a commercial bank for a commercial bank deposit, as private individuals do not normally bank with a central bank. When this happens the reserves (central bank deposits) of the commercial banks and commercial bank deposits are increased by the amount of the purchase. The reserve ratio of the commercial banks, however, has risen, and they are now in a position to expand their asset holdings and deposits up to the point where the reserve ratio once more is at the minimum level.

Regulation of Member Bank Reserve Ratios. Another possible weapon of central bank policy is the regulation of the reserve ratio of the member banks. A rise in the legal reserve ratio automatically curtails the maximum amount which member banks can lend, and a fall automatically raises that amount.

Banks Have More Control Over Their Assets Than Over Their Liabilities

It should now be clear that the most active part of banking is making loans. A banker does not have much control over the amount of deposits which he gets, although up to a point he can encourage depositors by offering higher rates of interest or greater conveniences, or by advertising. This is done to some extent, but on the whole banks do not go in for the competitive attraction of depositors. The rates of interest on deposits, even on savings deposits, are generally fixed by custom or agreement, as are the charges for checks, etc. In the making of loans, however, a banker has a good deal of discretion. He can attract or discourage loans in two ways: either by manipulating the rate of interest which he charges or by manipulating the other conditions of a loan. Generally speaking, the lower the rate of interest, the more applications for a loan will the banker receive, and the higher the rate of interest, the fewer applications will he receive. A banker is not obliged to grant a loan to anyone who applies, however; in this he differs sharply from a shopkeeper, who usually sells to anyone who comes to buy. A banker has to look to

the repayment of the loan; and if he feels that the risk of nonpayment is too great, he will not grant the loan. Or, if he feels that he does not wish to expand his loans, he can still refuse to grant the loan. even if the person seeking the loan is in sound financial standing.

The "Rationing" of Loans

This power of the banker to make loans or not, as he sees fit, is of great importance. We may call it his power of "rationing" loans. As we saw in Volume I, Chapter 11, the price of any commodity can be kept permanently below its equilibrium price only if there is rationing of some kind. If a price is below its equilibrium level, buyers will wish to buy more than sellers wish to sell. If buyers are to be prevented from raising their bids, then, there must be some way of apportioning among them the small quantity which is available. If a rise in price to choke off the buyers is not allowed, some other way of choking them off must be found. This other way is rationing. In the case of loans, then, the rate of interest is an index of the "price" of loans. If at a given rate of interest borrowers wish to borrow more than the banker wishes to lend, in a competitive market the result will be a rise in the rate of interest. The market for bank loans, however, is not a competitive market. If at a given rate of interest borrowers wish to borrow more than a banker wishes to lend, the banker does not have to raise the rate of interest in order to choke off the too-eager borrowers. All he needs to do is put on a stern face and refuse the borrower the loan when he applies. This fact is of great importance, especially in interpreting banking history.

Bank Purchase and Sale of Securities

Besides making loans, banks can also buy and sell securities, usually government securities, and this also is a fairly active part of banking. The purchase of securities by a bank has much the same effect on the expansion of deposits as the making of a loan; indeed, the making of a loan is itself merely the purchase of a special kind of security—a promissory note. The effect on the overall economy, however, is very different; bank loans are usually made to finance new investment, whereas when a bank purchases securities, this merely withdraws securities from other owners and has little direct impact on investment except in so far as the sellers of the securities may use the bank deposits acquired for new production.

Central Bank Policy

The degree of control which a central bank is able to exercise over the monetary system depends on the nature of the reactions of commercial

banks and other institutions to its policies, and also on the objectives of these policies themselves. Central banks operate directly by withdrawing or releasing securities of various kinds, thus affecting the holdings of these securities in the hands of other persons or institutions, and also by creating or withdrawing central bank deposits. The effects of these operations depend of course on the stability of the reactions of money market agencies and organizations. If the central bank contracts its deposits by selling securities, the commercial banks may be able to offset this contraction in their reserves by lowering their reserve ratios, or by other devices designed to economize reserves. In the depression of the 1930s the expansionary policy of the Federal Reserve System was, in part, offset by a rise in the "voluntary" reserve ratio of the commercial banks—i.e., a rise in their reserve ratio above the legal minimum. In the inflation of the 1940s the ability of the Federal Reserve System to offset the inflationary effects of war finance and its aftermath was largely destroyed because of the policy of supporting the price of government bonds. Between about 1942 and March, 1951, the Federal Reserve System virtually established a "government bond standard," by standing ready at all times to buy government bonds if their prices fell below a level which would yield an appropriate rate of interest. This meant that all government bonds which the market did not want to hold at the established rates of interest found their way into the portfolios of the Federal Reserve banks, with consequent enlargement of Federal Reserve deposits and hence of member bank reserves.

Effects of a Banking System

It is clear that the development of a banking system produces great effects on the economy, both on the volume of securities of various kinds, the volume of liquid assets, and also on the price level and on the volume of payments and of incomes. It is impossible at this stage to follow all these ramifications. Simple models of the effect of the introduction of a banking system can be developed with the aid of an extended form of the Fisher identity (page 77).

Let us suppose, for instance, a system in which there is no central bank, in which there is only one form of legal money (say gold coin), and in which bank reserves are all in the form of legal money. Then if M is the quantity of legal tender money in the possession of the public, and V is its velocity of circulation, the product MV is the total flow of money directed toward the purchase of commodities which are bought with legal money. Similarly, if D is the total quantity of bank deposits in the possession of the public, and V' is the velocity of circulation of these deposits,

the product DV' is the total flow of money directed toward the purchase of those things which are bought with bank deposits. As practically all commodities are bought either with legal money or with bank deposits, the total flow of money directed toward the purchase of commodities must be $MV + DV'$. This, as we saw in the previous chapter, is equal to the value of the commodities bought, or PQ. We can therefore rewrite the Fisher equation:

$$PQ \equiv MV + DV' \tag{11}$$

If there were other forms of money in the system, say, private bank notes, these could be brought into the equation in the same way. If the quantity of this form of money were D_1 and its velocity of circulation were V'', then the flow of money directed toward the purchases of goods bought with this kind of money would be D_1V'', and the equation would become:

$$PQ \equiv MV + DV' + D_1V'' \tag{12}$$

In a similar way it could be extended to include any number of kinds of money used in the purchase of commodities.

How Banking Increases Payments

We can now see how the banking system increases payments. Suppose, first, a system without a central bank. Then equations (1) to (8) hold whether the system has one or many banks, provided we interpret $\bar{r}$ to be the reserve ratio of the whole system; $\bar{h}$, to be the preferred ratio of cash held by the public to *all* bank deposits; D, to be the total of bank deposits; M, to be the total cash held by the public; and N to be the total cash. Then, as we see from equations (5) and (7), if E is the total expenditures before, and E', the total payments after, the introduction of banking

$$E = NV \tag{13}$$

$$E' = \frac{\bar{h}NV}{\bar{h}+\bar{r}} + \frac{NV'}{\bar{h}+\bar{r}} = \frac{E(k+\bar{h})}{(\bar{r}+\bar{h})} \tag{14}$$

where $\qquad V' = kV$

That is to say, the result of the introduction of banking will be to increase the total volume of payments provided $k > r$. As k is usually greater than 1, and r about 0.1, it would take a most unusual situation to prevent the introduction of banking from substantially increasing the total volume of payments. To give some figures for illustration, if $r = 0.1$ and $h = 0.3$ and $k = 2$, then the introduction of banking will increase the volume of payments by nearly six times.

Effect of Central Banking. The introduction of central banking changes the model somewhat. The reserves of the member banks are now determined mainly by the volume of assets held by the central bank. If, for instance, previously the reserves of the commercial banks were in the form of gold, the introduction of a central bank and the centralization of the gold stock into the possession of the central bank may increase the total of the commercial bank reserves, as the central bank may hold other assets besides gold. In a central banking system also the decisions of the public as to the desired proportion of bank deposits to other forms of money can largely be neutralized by appropriate central bank action. Thus, if the American public suddenly decides to hold a larger proportion of "folding money" (federal reserve notes, for the most part) to bank deposits, the Federal Reserve System stands ready to expand its note issue to accommodate the new demand, and at the same time to expand its holdings of securities enough to prevent a serious decline in member bank reserves and deposits.

Events Table for a Central Banking System

These principles are illustrated by a simple events table for a central banking system, as shown in Table 11. The assets of the central bank are assumed to consist of only gold, G_c, in line 1 and securities, S_c (usually government securities), in line 2. The liabilities of the central bank are notes, M (line 4), which constitute the pocket cash (folding money) of the public, and deposits, R (line 5), which constitute the reserves of the commercial (member) banks (line 6). Line 7 is the securities, S (loans and investments), held by member banks; line 8 is the sum of their assets, $R + S = D + C$, where D equals deposits of member banks (line 9), and C equals capital (net worth) of member banks (line 10). Line 11 shows the total money stock, $D + M$; line 12, the reserve ratio, $r = R/D$; line 13, the preferred cash ratio of the public, $h = M/D$. We suppose that the velocity of circulation of M is 2 and of D is 6, so that in line 16 we have the total volume of payments, $2M + 6D$.

The following are the equations of the system. We have first the two balance-sheet identities:
For the central bank

$$G_c + S_c = M + R \tag{15}$$

For the member banks

$$R + S = D + C \tag{16}$$

Then we have the two behavioral equilibrium conditions of equations (3) and (4), p. 103. The system is the same as equations (1)–(8), except

TABLE 11. A CENTRAL BANKING SYSTEM

	Period								
	1	2	3	4a	5a	6a	4b	5b	5c
Central bank									
Assets									
1. Gold (G_c)	—	2,000	2,000	2,000	2,000	2,000	1,500	1,500	1,500
2. Securities (S_c)	—	—	1,000	1,000	1,482	1,640	1,000	1,500	1,000
3. Total	—	2,000	3,000	3,000	3,482	3,640	2,500	3,000	2,500
Liabilities									
4. Notes (M)	1,500[a]	1,500	2,250	2,400	2,786	2,912	1,875	2,250	2,250
5. Deposits (R)	—	500	750	600	696	728	625	750	250
Member banks									
Assets									
6. Reserves (R)	500[a]	500	750	600	696	728	625	750	250
7. Securities (S)	5,500	5,500	7,750	6,400	7,268	7,551	6,625	7,750	8,250
8. Total	6,000	6,000	8,500	7,000	7,964	8,279	7,250	8,500	8,500
Liabilities									
9. Deposits (D)	5,000	5,000	7,500	6,000	6,964	7,279	6,250	7,500	7,500
10. Capital (C)	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
11. Money stock ($D + M$)	6,500	6,500	9,750	8,400	9,750	10,191	8,125	9,750	9,750
12. $r = R/D$	.1	.1	.1	.1	.1	.1	.1	.1	.033
13. $h = M/D$	.3	.3	.3	.4	.4	.4	.3	.3	.3
14. $2M$	3,000	3,000	4,500	4,800	5,572	5,824	3,750	4,500	4,500
15. $6D$	30,000	30,000	45,000	36,000	41,784	43,674	37,500	45,000	45,000
16. $E = 2M + 6D$	33,000	33,000	49,500	40,800	47,356	49,498	41,250	49,500	49,500

[a] Gold.

that $N = G_c + S_c$. If then we are given $G_c + S_c$, this determines D by the equation, like (5).

$$D = \frac{G_c + S_c}{\bar{h} + \bar{r}} \tag{17}$$

Then given $\bar{r}$ and $\bar{h}$, we have $R = \bar{r}D$ and $M = \bar{h}D$. C is given, so $S = D + C - R$.

Effects of Introducing Central Banking

Suppose, then, we start in column 1 before the introduction of central banking; this is the situation of column 6, Table 10, p. 103. M here represents gold in the hands of the public as cash; R, gold in the reserves of commercial banks. With the introduction of central banking in column 2, the gold is centralized in the central bank, G_c, and the public is given notes and the member banks, central bank deposits in exchange; the other variables are unchanged. Now, however, in column 3 the central bank buys 1000 of securities with either notes or deposits. If the public is free to exchange its member bank deposits, D, for central bank notes, M, this process will go on until the ratio h comes back to its preferred value, assumed to be 0.3. The result, as we see by comparing columns 2 and 3, is a considerable expansion both of the total money stock (line 11) and the total volume of payments (line 16).

Effects of Increase in Preferred Cash Ratio

In column 4a we see the result of a spontaneous increase in the preferred cash ratio from .3 to .4 without any compensatory action on the part of the central bank. The public buys central bank notes, N, with some of its deposits, D, which forces the member banks to exchange some reserves, R, for notes, M; the decline in their reserves forces them to contract their own loans and investments, S, which further reduces their deposits, until the system shakes down to the position of column 4a. There is quite a sharp deflation, reflected both in a fall in the total stock of money (line 11) and in payments (line 16).

Suppose now that the central bank decides to compensate for this deflationary move by expanding its security holdings. It can restore the total money stock of column 3 (9,750) in column 5a by expanding its holdings of securities to 1482.[2] If it wishes to restore the total volume of payments of column 3 (49,500), it must expand its security holdings further to about 1,640.[3]

[2] If $D + M = 9{,}750$ and $M = 0.4D$, $1.4D$, $= 9{,}750$ and $D = 6{,}964$; hence $M = 2{,}786$ and $R = 696$; then $S_c = M + R - G_c = 1{,}482$, and $S = D + C - R = 7{,}268$.

[3] If $2M + 6D = 49{,}500$ and $M = 0.4D$, $6.8D = 49{,}500$, $D = 7{,}279$; wherefore $M = 2{,}912$, $R = 728$, $S_c = M + R - G_c = 1{,}640$, and $S = D + C - R = 7{,}551$.

Effects of a Drain of Gold

In column 4b we see the result of an uncompensated drain of gold; comparing columns 4b and 3 we see that the central bank has "sold" 500 of gold, for its notes or deposits, reducing its total assets to 2,500. It does not matter whether the initial transaction consisted of exchanging notes or central bank deposits for gold; if notes are freely available for deposits, the public will rearrange its holdings until the situation of column 4b obtains. We see there is a sharp deflation, with both money stock and total payments declining sharply. To offset this (column 5b), the central bank must expand its holdings of securities so that the old asset total is reached, for the other variables of the system depend only on the total $G_c + S_c$. It may do this partly by lowering its rediscount rate, which will induce member banks to borrow from the central bank, or it may employ the more powerful method of purchase of securities in the open market (open market policy). If the central bank itself had to keep a constant reserve ratio, say $(M + R)/G_c$, its power to offset gold movements by changes in its own asset total would be virtually nil. However, if it had the power to change the reserve ratios of the member banks, it might still have a powerful offsetting instrument. This is shown in column 5c, following 4b, where the member banks' reserve ratio is reduced to .033, which restores the money stock and payments of column 3.

EFFECTS ON VOLUME OF TRANSACTIONS AND OUTPUT. If the output of goods in a society remained the same, we should expect the introduction or extension of banking to result in a large rise in prices. Historically, however, the development of banking has gone hand in hand with a large increase in the total output of society, an increase due partly to increasing population, partly to improved techniques of production. It is true that the price level has increased considerably since the seventeenth century, which saw the beginnings of large-scale banking in the Western world. Nevertheless, the main result of the development of banking has been to prevent the great *fall* in prices which would otherwise have resulted from the expansion in population and output. How far the absence of banking would have prevented this expansion of output is an interesting problem to which there seems to be no definite answer. It is true that a rising price level encourages enterprise, even while it may often guide it into wrong channels. It is true also that an expansion of bank loans implies that there is an increased quantity of capital goods, for loans are issued mainly for the purpose of enabling entrepreneurs to create goods. An expansion of bank loans, therefore, unless the loans are sadly misapplied, should result from an expansion of the volume of transactions and of output. This will have the effect of moderating

the fall in prices which would otherwise occur, as will be seen from the Fisher identity. It can be argued that there is some rate of expansion of the banking structure which will just enable the flow of money to keep pace with the rising volume of transactions, thereby keeping the price level constant. Some writers believe this to be an ideal of banking policy. This question, however, must be reserved for a later discussion.

APPENDIX

THE MANY-BANK SYSTEM

Suppose we have n commercial banks in a system without a central bank. Let the i^{th} bank have deposits D_i, reserve R_i, capital C_i, and securities S_i, and a reserve ratio $r_i = \frac{R_i}{D_i}$. Let N be the total quantity of cash (reserve) money, and M, the quantity in the hands of the public, so that

$$N = M + R_1 + R_2 + \dots + R_n \tag{18}$$

Let total deposits $= D$, total reserves $= R$, total securities held by banks $= S$, total bank capital $= C$. Then let us postulate a behavior ratio, $h = \frac{M}{D}$, as before, and an equilibrium value of this h, the preferred cash ratio, such that if $h < \bar{h}$, people will, on balance, deposit cash, and if $h > \bar{h}$, people will withdraw cash from the banks. Then let us postulate for each bank, such as the i^{th} bank, a ratio $b_i = \frac{D_i}{D}$, with an equilibrium value, $\bar{b}_i$, representing the relative demand for the deposits of the i^{th} bank, such that if $b_1 < \bar{b}_i$, people will transfer deposits to the i^{th} bank,

and if $b_1 > {}_i\bar{b}_i$, people will transfer deposits from the i^{th} bank to others. We then have

$$D = \frac{D_1}{b_1} = \frac{D_2}{b_2} = \frac{D_i}{b_i} = \frac{D_n}{b_n} = \frac{M}{h} \tag{19}$$

From (18) and (19) we have

$$N = D_i\left(\frac{h}{b_i} + \frac{r_1 b_1}{b_i} + - + \frac{r_n b_n}{b_i}\right)$$

That is

$$D_i = \frac{N b_i}{h + r_1 b_1 + \ldots + r_n b_n} \tag{20}$$

$$= \frac{N b_i}{h + r}$$

where r is the average reserve ratio, $\frac{R}{D}$

We also have

$$D = \Sigma D_i = \frac{N}{h + r}, \text{ as } \Sigma b_i = 1 \tag{21}$$

Equations (20) and (21) should be compared with equation (5), p. 104. In equilibrium

$$D_i = \frac{N\bar{b}_i}{\bar{h} + \bar{r}} \tag{22}$$

that is, the volume of deposits of a particular bank depends on the strength of the preference for that bank relative to others, and on the overall reserve ratio of all banks and the overall preferred cash ratio. We may observe that D_i does not depend directly on r_i, except in so far as a fall in r_i will change r; that is, a bank which has a low reserve ratio does not thereby increase its deposits, for it loses reserves to others; indeed, if the low reserve ratio lessens confidence in and preference for the bank, the bank may actually lose deposits.

QUESTIONS AND EXERCISES

1. Suppose we have a system with a single bank. What will be the immediate and ultimate effects on (a) the assets, liabilities, and reserves of The Bank, and (b) the assets and liabilities of private persons, of the following actions?
   a. Mr. Smith takes \$100 in cash and pays it in to The Bank.
   b. Mr. Smith pays in a check for \$50, received from Mr. Jones.
   c. The Bank discounts a three-months' bill for \$5000 for Mr. Robinson.
   d. Mr. Smith asks for and receives \$100 of his deposit in the form of bank notes.

2. Suppose we have a system with two banks, one of which has ten times the volume of deposits, and ten times the number of depositors, as the other. What will be the probable effect on both banks (assets, liabilities, and reserves) of the repayment by Mr. Jones, in cash, of a loan of $5000 at the larger bank? Be clear as to your assumptions.
3. Continue the events table of Table 10 with the following successive events:
   a. The Bank lowers its reserve ratio to .08;
   b. the preferred cash ratio of the public falls to .25;
   c. a loan of 500 is defaulted and The Bank collects only 200;
   d. N decreases by 500, due, say to an outflow of gold;
   e. suppose The Bank cannot call in any more loans, but that h rises to the point where there is a "run." At what value of h will The Bank suspend payments, that is, run out of reserves?
4. It has been proposed that banks should be compelled by law to have a 100 percent reserve ratio. Discuss the possible effects of such a law on the practices of banks and on the power of the banking system to affect prices.
5. Suppose there was a system which possessed $600 million of legal money, in which the income velocity of circulation of legal money was 25 times a year, and of bank deposits, 50 times a year. Suppose the banks kept a reserve ratio of 20 percent, and that people wished to have $7 of bank deposits for each $1 of legal money. What would be the total money receipts of such a society, assuming that all bank reserves consist of legal money?
6. In Table 11, page 115, suppose that the central bank followed the policy of holding a constant ratio of its deposits plus notes to its gold holdings, so that $M + R = 2G_c$. What would be the equilibrium position of the system, the other assumptions being held constant as in column 3?
7. Continue the events table of Table 11 for the following successive events, starting from column 3:
   a. suppose the velocity of circulation of notes rises to 2.5 and of deposits to 7 times a year, but no change in policy is recorded;
   b. the central bank now acts to restore the old volume of payments;
   c. because of a war the banks (both central and member) have to increase their holdings of securities by 5000.
8. Adapt Figure 11 to the model of Table 11 for a central banking system.
9. The banks of the United States hold a considerable proportion of their assets in the form of government securities. What then, after World War II, would have been the effect of a sharp fall in the price of government securities on the solvency of banks? Might this effect be expected to hamper the Federal Reserve System in its open market policy? If so, how?
10. Suppose the government raised the rate of interest on Post Office Savings certificates from 2 to 3 percent. What effect might this have on the banking system?
11. "Banks do not create money; they merely increase its velocity of circulation." Discuss critically.
12. Explain, by means of the Fisher identity, how speculation may affect prices.
13. In the Appendix, develop formulas for R_t, S_t, R, and S.

CHAPTER 6

SAVING AND INTEREST

THE CONCEPT OF SAVING

With the tools of analysis developed so far we can throw a good deal of light on two dark areas of economic theory—the concept of saving and the theory of interest.

Difficulties in the Concept of Saving

The concept of saving has caused a great deal of trouble in economic theory, and even though it can be dispensed with almost entirely in modern economics, its traditional importance is such that it may be well to relate it to the concepts of the preceding chapters. The difficulties created by the concept have been of two kinds. One has arisen because saving is equated in some sense with the idea of "not consuming," and negative concepts of this kind always cause difficulty: the realm of the might-have-beens is always difficult to define. Thus it is much easier to say what one had for dinner last night than to say what one did not have for dinner; and if saving is not consuming, the question, What did you not consume? becomes almost meaningless. For this reason it is generally much safer to use the consumption concept than the saving concept; consumption as an *activity* is much more positive and easy to define than saving. The other difficulty has arisen because of the confusion between saving in the sense of hoarding (accumulating liquid assets) and saving in the more proper sense of the accumulation of net worth, or assets in general. This is an example of a common confusion between the process of production, consumption, and accumulation, on the one hand, and the process of the circulation of money by means of "payments," on the other.

The Savings-Investment Controversy

A good example of the difficulties involved in the concept of saving is to be found in the so-called "savings-investment controversy" which is associated with the work of J. M. Keynes. As early as 1924, Keynes[1] hinted that economic fluctuations might be a result of the fact that investment was performed by one set of people and saving by another set, and that there was no direct relationship between these two sets of activities. In his *Treatise on Money* in 1931 Keynes tried to define saving and investment in such a way that the divergence between them could be used in the explanation of monetary changes. This attempt, however, was unsuccessful, and in his *General Theory of Employment, Interest and Money* (1936) Keynes finally formulated the *identity* between savings and investment which is the basic identity of modern macroeconomics. This identity in real terms is nothing more or less than the basic identity (equation (4) on page 49).[2]

Accumulation = production minus consumption = (saving)

or
$$A = P - C \tag{1}$$

Production means adding to the total stock; consumption means subtracting from the total stock; investment, in real terms, is the net amount added to the total stock of real capital. Production is real income; consumption is real "outgo." Saving in real terms is the excess of real income over consumption. In real terms, therefore, investment and saving are identical, in the sense that they are simply different ways of looking at the same phenomenon.

Another aspect of the savings-investment identity is the "sources and uses of savings" identity as we find it in the national income accounts, either in the simplified form of equation (7), p. 35, or in the more elaborate form of equation (24), p. 44.

Effect of Price Changes on Saving

A difficulty arises when we are considering saving and investment in their financial, or "dollar value," aspect. Investment in this sense means the growth in the value of real assets. The value of real assets, however, can grow not merely because additions to real assets (production) exceed

[1] J. M. Keynes, *A Tract on Monetary Reform*, New York, Harcourt, Brace, 1924

[2] This is the proposition which in a somewhat frivolous mood I have labeled the "bathtub theorem"—the total stock being the water in the tub; production, the flow from the faucet; consumption, the flow down the drain; the difference between production and consumption being the rate of accumulation of water in the tub. This is positive if the inflow (production) exceeds outflow (consumption)—negative if outflow exceeds inflow.

subtractions (consumption) but also because the prices of existing assets rise. Suppose, to take a simple example, that there is only one real asset, wheat. Suppose, now, that the stock of wheat at the beginning of a year is 500 million bushels, and that 800 million bushels are produced and 700 million bushels consumed during the year. Evidently, the stock of wheat will rise by 100 million bushels to 600 million bushels during the year. If the price of wheat were constant at $2 per bushel, the value of the wheat at the beginning would be $1000 million and at the end, would be $1200 million, the rise of $200 million (investment) representing exactly the difference between income ($1600 million per annum) and outgo ($1400 million per annum). If, however, the price of wheat changes during the year, the situation is more complicated. Suppose, for instance (to avoid difficulties in the valuation of income and outgo), the price remains at $2 per bushel all through the year and jumps suddenly to $3 on the last day of the year. The value of the 600 million bushel stock is now not $1200 but $1800 million—the value of the stock has risen during the year not by $200 million but by $800 million, whereas the difference between income and outgo remains at $200 million. What are we to do in such a case? Two alternatives are open—we can either redefine income to include changes in the value of an existing stock due to a change in its price (capital gains) or we can restate our original proposition in the form,

$$\text{Investment} = \text{saving} + \text{capital gains}$$

These alternatives are of course merely verbal; in both cases the full equation reads

$$\text{Increase in value of stock} = \begin{array}{c}\text{value of production} - \text{value of consumption} \\ + \text{ capital gains} - \text{capital losses}\end{array}$$

In the first alternative we define income as the value of production plus net capital gains; in the second, we leave the definition of income as the value of production alone. The ambiguity in the concept of income revealed by this analysis shows up in the United States income tax, according to which some capital gains are recognized as income but are taxed at a different rate from other income, thus paying tribute to the fact that capital gains are in some sense income but of a peculiar kind!

Hoarding

In the case of an individual the term "saving" sometimes carries the meaning of the excess of *receipts* (money payments in) over *expenditures* (money payments out). The term "hoarding" is sometimes used to express this quantity; it is of necessity equal to the increase in money

assets. The difference between saving (the increase in total assets) and hoarding (the increase in money assets) may be seen most clearly perhaps from an example. Suppose a man has at the beginning of a year a total net worth of $20,000, of which $2,000 is in the form of money and $18,000, in the form of goods. Suppose during the year he produces (earns) $10,000—i.e., he adds $10,000 worth to his total assets—and suppose that $8,000 worth of assets are consumed. Clearly, his total assets (neglecting capital gains) are now $22,000, and he has saved $2,000. While he has been saving $2,000 in total assets, suppose that his total money receipts have been $12,000. His receipts exceed his income, indicating that he has received money not only from the sale of his current production or labor but also from the sale of some previously possessed asset—say, a piece of land. His money expenditures, we suppose, have been $11,000. These exceed his consumption, indicating that his purchases were not all immediately consumed, but that some of his purchases were in the form of durable goods—say, a house or an automobile. We can immediately deduce that his money assets will have increased by the difference between money paid in and money paid out—i.e., $12,000 – $11,000 or $1,000. At the end of the year, therefore, his money assets must be $3,000 and his non-money assets $19,000. He has saved $2,000, but has only hoarded $1,000. It is evident from this example that it is quite possible for an individual to be saving and dishoarding—i.e., increasing his total assets but diminishing his money assets—in the same period. It is equally possible to hoard and dissave simultaneously.

Keynes' Concept of Saving and Investment

It should be pointed out that the concepts of saving and investment in Keynes' *General Theory* are not quite the same as the concepts developed above. Keynes defines saving as the excess of income over *consumers' expenditure,* not as the excess of income (production) over consumption. Consumers' expenditure is equal to the value of consumer (i.e., household) purchases. It differs from consumption because many household purchases are of durable goods which are not immediately consumed, and consumption, both of consumer durables and of goods held by businesses, is a process quite different from the processes of purchase and sale, even though the two processes are often closely associated.

In the United States national income accounts, for example, in the simplified form of equation (6) p. 35, or in the form of equation (25) p. 44, saving is defined in Keynesian terms as the difference between disposable personal income and consumers expenditure. This underestimates, by a considerable sum, the amount of saving and investment in terms of real accumulation, as it neglects the fact that a good deal of house-

hold expenditure represents the purchase of durable goods, and for most households the value of durable goods purchased exceeds the depreciation of old household goods. Thus, if we did national income accounts in terms of real accumulation we should divide personal consumption expenditures, B_1, into two parts: personal consumption B_{1c}, and personal accumulation, B_{1a}. Then we should add personal accumulation to personal saving, C_7, to total saving, B_{18}, and to total investment, H_{18}, as defined in the national income accounts to get the savings-investment identity in terms of total accumulation. The national income accounts do break down personal consumption expenditures (Table B-2 in the *Economic Report of the President,* January 1965) into durable goods, C_2, nondurable goods, D_2, and Services, E_2. Durable goods expenditure by households was $51.9 billion in 1963, by comparison with $83.3 billion for gross private domestic investment, so we see that household accumulation is at least in the same order of magnitude as business accumulation. Unfortunately, we do not have any estimate of household capital consumption, so we do not have a figure for net household investment.

Saving Not "Postponed Consumption": Satisfaction Derived from Use, Not Consumption

It is now necessary to discuss yet another aspect of the saving concept, the notion of saving as "postponed consumption." This is common in economic literature, yet it seems to embody a profound misconception of the nature both of saving and of consumption. The difficulty seems to arise because of a confusion between the idea of consumption as we have used it in this work, and as it is also used by the classical economists, as the "using up" or the destruction of stocks of goods, and the idea of consumption as the source of satisfactions or enjoyments. In fact, these are two quite distinct ideas. Enjoyment or satisfaction is normally derived not from the using up but from the *use* of a good. This is obvious in the case of durable goods. When I go for a ride in my car, the fact that the car is being consumed—i.e., is wearing out—in no way contributes to my satisfactions; indeed, quite the reverse. What I get satisfaction from is the use of the car, not the fact that rust, wear, and grit are nibbling away at it. If I had an indestructible car, my satisfaction in the ride would in no sense be diminished. Similarly, if we had unbreakable china, clothes that never got soiled or worn, houses that never decayed, and machines that never wore out, consumption would be much diminished, but we would be all the richer for that. Satisfactions, therefore, must be regarded as proceeding from the stock of goods, not from their consumption. It is the size of the house we live in, the elegance of the clothes we wear that give us satisfactions, not the incidental and regrettable fact that in the course

of yielding these satisfactions, or even when they are not yielding satisfactions, they happen to be afflicted by decay and consumption.

Economy in Consumption Always Desirable

The reader may object that this argument does not apply to nondurables, or "one-use goods," such as food, which are destroyed in the very act of enjoying them. It is only because consumption and satisfaction are close together in time, however, that we tend to identify them; analytically they must be kept separate. Even in the case of one-use goods, economy in consumption is always desirable, and if we examine the use of these goods closely we will always find that they are consumed in the maintenance of some desired but depreciating state. Thus we burn fuel because the warmth of our houses "depreciates" in a cold climate. The *less* fuel we can burn to maintain a given temperature, the better off we are. We need food similarly to maintain certain bodily states which likewise depreciate. The less food we need to keep us from being hungry and to maintain our bodily temperature and energy, the better off we are. Similarly we need entertainment to maintain certain mental or emotional states (see p. 16). It should be observed that the depreciation of goods (i.e., of desirable states) frequently involves their *physical* growth. Thus a shaved chin, a clipped head, or a mowed lawn depreciates by the process of growth of a "discommodity" (whiskers, hair, grass) and needs to be restored by periodic shaving, haircutting, or mowing. In this case it is the growth of the discommodity which constitutes consumption in the economic sense, and its removal which constitutes production.

Saving as Accumulation and as the Sacrifice of Consumption

In the light of these considerations, then, what is the significance of saving? Saving is the process by which goods are accumulated—i.e., by which the total stock of goods is increased. This can only be done by producing more than is consumed. If society is producing at capacity, saving does imply a certain sacrifice of consumption, in the sense that consumption would be greater if there were no saving. The sacrifice of consumption is also likely to lead to a sacrifice of satisfactions, unless there is at the same time an economizing of consumption—i.e., a decline in the amount of consumption which is necessary to maintain given states of satisfaction. Because of these sacrifices, however, the total stock of goods (desirable states) is increased, and the total *future* flow of satisfactions is presumably increased also. In this sense saving does involve the sacrifice of present satisfaction in order to increase satisfaction in the future. It does not, however, necessarily involve the sacrifice of present consumption in order to increase future consumption. It is true that

future consumption may be increased as a result of saving, both because a larger stock of goods in itself implies a higher rate of consumption and because the accumulation of the larger stock may permit higher rates of production. We do not generally, however, accumulate now *in order that* we may decumulate in the future; we do not build up a stock of goods in order to tear it down in the future, but in order to be able to maintain and increase this stock in the future. We have not built up all this apparatus of houses, farms, roads, machines, harbors, and factories in order that some day we may allow them to fall down and so return to living in holes in the ground and to grubbing our food from berry bushes!

Saving in a Subsistence Society

An illustration may clarify these principles. Suppose that we have a society which is living just above the edge of subsistence, with its productive activities just sufficient to maintain the bodily strength of its people, to reproduce the generations as they die off, and to maintain the miserable huts in which the people live and the scanty clothes which they wear. How does such a society ever progress to a better state? If it is to improve, it is clear that it must build up its stock of capital: it must have more implements, more machines, more livestock, better houses, and so on. In order to do this (without aid from outside) it must withdraw resources from maintaining the existing fabric of the society in order to devote them to making the increased stock of implements, etc. This means that some of its existing states cannot be maintained as well as before: people may have to go a little hungry in order that the implement makers can be spared from food-production; leisure time and ceremonial activities may have to be skimped in order to release time for building, and so on. All this involves curtailment of satisfactions as well as of consumption. If, now, the result of this process is simply a larger stock of things—better houses, better clothes, and so on—without any improvement in the productivity of the society, as measured in some sense by output per man-hour, the society may be no better off than before. In order to maintain its bigger houses and finer clothes the society may still have to withdraw resources from previously enjoyed occupations, though not so much as in the period of building up the stock. If, however, the increased stock is at least partly in the form of instruments and implements which increase output per man-hour, the society is *permanently* richer as a result of the accumulation; and if the improvement is sufficient, it will not only be able to maintain its increased stock with no more effort than it previously took to maintain its smaller stock, but may even be able to maintain the larger stock with less effort than it took to maintain the smaller

one. In that happy event the society will not only be richer, but will find it easier to get still richer, as it will be able to devote to further accumulation the resources released from maintenance by its increased productivity.

Individual and Aggregate Saving

Up to this point we have discussed saving in the aggregate; we have not examined adequately the relation between individual saving and aggregate saving. An individual saves, as we have seen, when he increases his net worth. The sum of individual net worths is equal to the total value of the real capital of society, so that when a society saves (increases the value of its real capital) individuals also must have saved (increased their net worths). Thus if the real capital of society increases in value by $1 million, this $1 million must show up as an increase in individual net worths. We may say perhaps that an increase in physical capital represents "real" saving, whereas an increase in the value of a constant physical capital, because of price increases, or an increase in government debt, is "false" saving. Real or false, however, any increase in the dollar value of real assets will show up as individual dollar saving, assuming, of course, that all debts cancel out when balance sheets are aggregated. It does not follow, however, that action on the part of a single individual to increase his own net worth, even successful action, will necessarily increase aggregate saving by an equal amount, because an individual may increase his own net worth at the expense of others, or even to the benefit of others. The most obvious case of this phenomenon is plain theft, where ownership of physical capital (say a diamond ring) is simply transferred from the robbed to the robber. There are many other cases of such transfer, however, many of them legal. In any speculative boom and crash, for instance, there are widespread redistributions of net worth, and the gains of one may be offset by the losses of another. An even more significant case is where the attempt on the part of many or most individuals in a society to increase their net worths by restricting their personal consumption results, not in accumulation at all, but in a diminution of production through unemployment.

The "Paradox of Thrift"

This is the famous "paradox of thrift" which has haunted economics in the underworld of Mandeville and Malthus,[3] and finally achieved respectability in Keynes. The restriction of consumption only leads to a corresponding amount of accumulation if production remains unchanged. Thus if production remains at 100, and consumption falls from 80 to 70,

[3] B. de Mandeville, *The Fable of the Bees,* London, 1714; T. R. Malthus, *Principles of Political Economy,* 2nd ed., 1836.

accumulation rises from 20 to 30. But if the decline in consumption from 80 to 70 results in a decline in production from 100 to 90, the decline in consumption produces no increase in accumulation. The "abstinence" in this case is not transmuted into "saving." And as we have seen in the "basic model" (pages 53–59), a fall in the consumption *function,* other things being equal—that is, a decline in the amount that will be consumed at all potential levels of income—will have the effect of decreasing the equilibrium level of output at which there is no excess of accumulation.

Saving and Age

Another factor of great importance in considering the relation between individual and aggregate saving is the age distribution of the population. There is a fairly regular pattern of accumulation and decumulation in the course of a single human life. In childhood and youth the individual generally consumes more than he produces. We do not, of course, carry the concepts of accounting in any exact way into the economic life of the family, but it is clear that in a sense the growth of children and young people represents a great deal of consumption which is not offset by any production of goods by them, but is offset in part by the growth of "human capital." Thus the young person entering his first job represents a substantial investment on the part of society. In the middle years the individual generally produces much more than he individually consumes. Part of the excess of his income over his personal consumption may be consumed in raising a family; part of it may be represented by a growth in his net worth—i.e., by saving. In old age the individual again returns to a condition where he is decumulating—i.e., where he is consuming more than he produces, and his net worth is diminishing.

"Rainy Day" Saving Does Not Result in Aggregate Saving

If the object of individual saving is simply to save for old age or for a rainy day, the individual over his whole life will not make any contribution to *aggregate* saving, as what he saves in his productive years will all be used up in his unproductive years. Similarly, a society in which all saving is of this nature will not achieve any net accumulation: the excess of income over consumption of those in middle life will just be offset by the excess of consumption over production of the young and the old.

Aggregate Saving and the Age Distribution

It should be observed, however, that the *age distribution* of the population has an important effect on the net result of "rainy day" saving. If

the population is heavily concentrated in the middle age groups it is quite likely to have an excess of attempted saving, or an underconsumption problem, because it is these age groups which consume less than they produce. If, on the other hand, the population is heavily concentrated in the young and the old age groups, it may run into an overconsumption problem—the consumption of the unproductive young and old constantly eats up all the excess production of the middle-age groups. These find it difficult to produce enough to support their dependents, and so the society finds net accumulation difficult. The difficulties of the United States in the 1930s were closely related to the fact that at that time an abnormally large proportion of its population was of working age. The immigrants of the great wave of 1900–1914 were still largely in the labor force, declining birth rates had strikingly reduced the proportion of children in the population, and the rising expectation of life had not yet had its full impact in increasing the numbers of old people. It is little wonder that the consumption function was low and that the economy was subject to violent depression. By the 1950s, however, the situation had changed. The proportion of old people was increasing rapidly, and also the remarkable rise in the birth rate from 1940 on had greatly increased the numbers of children and young dependents. The rise in the level of the consumption function since 1930, which has contributed in no small degree to the resiliency of the American economy since 1945, is in large part due to the change in the age composition of the population.

By contrast, an ironic tragedy is now working itself out in the poor countries of the tropics, where the introduction of chemical insecticides and other public health measures led to a dramatic reduction in infant mortality in the years around 1950 without a corresponding decline in the birth rate. The result is an unprecedented distortion in the age distribution, to the point where in many of these countries more than half the population is now (1964) under the age of 15. This imposes a burden on them so great that consumption can easily outrun production, and capital and income, decline. This retrograde movement from poverty into destitution is one of the major tragedies of our time.

Saving and the Expectation of Life

The tragedy is all the greater because it is a short-run dynamic consequence of a sudden rise in the expectation of life which in itself is a great help to development in the long run. Thus, in a society with a stationary or slowly changing population, the proportion of a population of working age depends in considerable degree on the average age at death. Where the average age at death is low, a large proportion of the resources of the society must be devoted to the simple replacement of its human capital, and a large proportion of the population will consist

of children. Such a society consequently finds accumulation difficult—the consumption of its product by children is always pressing on the meager productive powers of the adults. The difference in this regard is striking between the underdeveloped areas, where the expectation of life generally is about 30 to 35, and the developed, where the expectation of life is generally about 60 to 65. For this reason investment in longevity (health services, better food production, and so on) may be of outstanding importance for an underdeveloped culture, provided that it is accompanied by a sharp fall in the birth rate.

THE THEORY OF INTEREST

We have now come to the point in the analysis where we may profitably take up another problem of great importance to economics—the theory of the rate of interest.

The Rate of Interest as a Rate of Growth

We must first consider what are the dimensions of a rate of interest—that is, what kind of a quantity it is. A rate of interest is not a price, nor a simple exchange ratio, but a *rate of growth.* Thus if something grows at a rate of 5 percent per annum, this means that it increases during a year by $\frac{1}{20}$ its size at the beginning of the year. A rate of growth of 5 percent per annum is the same as a rate of $\frac{5}{1200}$, or $\frac{1}{240}$ per month. That is to say, the dimensions of a rate of growth, are $\frac{1}{\text{time}}$; the reciprocal of a rate of growth is a length of time. Thus we may express the same fact about an investment by saying that it yields 5 percent (that is, $\frac{1}{20}$ per annum, or that it can be capitalized at 20 years' purchase. The rate of interest is commonly defined as the ratio of a steady and perpetual income to the capital sum which will purchase it. Thus if $100 will buy an income of $3 per year for ever, the rate of interest is then 3 percent per annum. This, however, is merely a special case of the more general definition. If $100 grows at 3 percent per annum for a year, it will amount to $103. The sum of $3 can then be taken from it as income, and the process can be repeated indefinitely.

The Rate of Profit as a Rate of Growth

There are two aspects of the rate of growth of capital which are significant in economics. One is the gross rate at which the value of the total assets of firms grows through the processes of purchase, production, and

sale. This may properly be called the *rate of profit*. The other is the gross rate at which the value of contractual obligations such as bonds grows. This is properly called the *rate of interest*. Thus, suppose I have $100 in a savings deposit at 2 percent per annum. This begins as an asset of $100 in my balance sheet, as a liability of $100 in the balance sheet of the bank. In a year's time it will be $100 (1.02), or $102; in two years' time, $100(1.02)^2$, or $104.04; in n years' time it will be $100(1.02)^n$

Growth and Withdrawals

We have defined both the rate of profit and the rate of interest as a *gross* rate of growth because the net rate of growth of a capital sum depends on the sums which accrue to the liabilities of the account for dividend or interest charges. Thus suppose we have a firm, the total value of assets of which is $1,000,000 at the beginning of a year. Suppose that during the year, by the process of buying things, transforming things, and selling things, the asset total rises to $1,040,000 without any withdrawals for dividends or their equivalent. The rate of profit is clearly 4 percent per annum. If now the firm declares a dividend of $30,000—or, if this is a single proprietorship or partnership, if the owner or owners allocate and pay out $30,000 for their own use—the net worth total falls to $1,010,000, as $30,000 in cash will be withdrawn from it without any substitute. The rate of profit is not, therefore, the actual rate at which asset totals are growing, but the rate at which these totals would grow in the absence of withdrawals on the owner's account. Similarly the rate of interest in a contractual investment is not the rate at which the asset or liability actually grows, but the rate at which it would grow if there were no interest *payments*. Thus, if I own a security which pays me $3 per annum and which cost $100, if the payments were not made to me but added to the value of the account, the capital sum would grow at a rate of 3 percent per annum.

Growth of Net Worth

When a firm has contractual liabilities (e.g., bonds or promissory notes), the *net worth* of the firm grows at a rate which is greater the greater is the excess of the rate of profit over the rate of interest. Thus, suppose a firm starts at the beginning of a year with assets totaling $1000, liabilities of $600, and a net worth of $400. Suppose that during the year assets rise to $1100 in the course of the firm's business, the rate of profit being 10 percent per annum. Suppose the firm incurs a 5 percent per annum interest charge on its liabilities. Before making the interest payment, the liabilities at the end of the year will amount to $600 (1.05), or $630, and the net worth will therefore be $1100 – $630, or $470. The net

worth has grown at a rate of $\frac{70}{400}$, or 17.5 percent per annum.[4] If the rate of interest had been, say, 3 percent, the liabilities would have grown to \$618 and the net worth to \$482 at a rate of $\frac{82}{400}$, or 20.5 percent per annum. It should be observed that the actual payment, as opposed to the mere accrual of interest, makes no difference to this result. If, in the first example, the firm paid \$30 interest at the end of the year, this would reduce the assets to \$1070 through a diminution of cash, and would reduce the liabilities to \$600; the net worth would remain at \$470. It is only the *accrual,* not the *payment* of interest, which reduces the net worth.

Rate of Interest Determined by the Price of Bonds

We must now prove an important proposition: that the rate of interest in any *given* set of contractual obligations is inversely related to, and is *determined by,* the present price of these obligations. A set of contractual obligations may be called a "bond," for want of a more general term. On the part of the issuer it consists of an obligation to pay certain benefits at definite dates in the future. From the point of view of the owner of the bond it is a right to receive these benefits. The benefits, and the dates, are specified or are understood in the contract. The rate of interest on the bond is that rate of growth or accrual of the initial value of the bond which will just permit all the designated benefits to be paid without there being any deficiency or excess of value at the end of the period. It is clear that the smaller the initial value, or price, of the bond the greater must be the rate of interest accrual if any *given* set of benefits is to be paid. That is, the smaller the present price of a given expectation, the greater the rate of interest in the investment.

This principle can be illustrated in a simple case: a promissory note in which John Doe promises to pay \$105 to the holder at a date one year

[4] If the rate of profit is r_a, the rate of interest, r_e, the rate of return on net worth (i.e., the rate of growth of net worth), r_n, A is the total value of assets at the beginning of the year, L the total value of liabilities, and N the net worth. Then at the end of the year we have as the balance sheet identity

$$N(1 + r_n) \equiv A(1 + r_a) - L(1 + r_e),$$

that is, as

$$N \equiv A - L,$$

$$r_n \equiv \frac{Ar_a - Lr_e}{A - L} \equiv \frac{A}{N}(r_a - r_e) + r_e$$

It should be observed that when $r_a = r_e$, $r_n = r_a = r_e$—that is, when the rate of interest is equal to the rate of profit—both rates are also equal to the rate of return on net worth.

r_n will be zero when $r_e = \frac{A}{L} r_a$.

from the present. Whoever owns this document now owns an expectation of receiving \$105 in one year. We shall suppose that there is no doubt about the fulfillment of the promise, so that the element of risk can be neglected. The rate of interest on such a security would be equal to the ratio of the *increase* in its value to the original value. If, for instance, Richard Roe pays \$100 for this note, he will in effect exchange \$100 of present money for \$105 in one year's time. The total *amount* of interest he will receive is (\$105 – 100), or \$5. The *rate* of interest his capital will earn is $\frac{5}{100}$ per annum, i.e., 5 percent per annum. If he paid \$95 for the note, he would be exchanging \$95 now for \$105 in one year's time. The amount of interest would be \$10 and the rate of interest would be $\frac{10}{95}$ per annum, or 10.5 percent per annum. Similarly, if he paid \$102 for the note, the rate of interest would be $\frac{3}{102}$, or 2.94 percent. If he paid \$105 for the note, the rate of interest would be 0 percent per annum. It is evident from these examples that the greater the price paid for a security representing a *given* expectation of future payments, the *smaller* is the rate of interest in the purchaser's investment. The principle is true no matter how many, or how frequent, the expected payments may be.

Consider another simple case—that of a "perpetuity," such as British Consols. The possession of one of these bonds entitles the owner to receive a stated sum, once every year, in perpetuity. Suppose we have a bond, for instance, which entitles us to receive \$50 on the first of January every year. If we paid \$1000 for this bond, the rate of interest on our perpetual investment would be $\frac{50}{1000}$ or 5 per cent per annum. If we paid \$2000 for the bond, the rate of interest would be $\frac{50}{2000}$, or 2½ percent per annum. If we paid \$500 for the bond, the rate of interest would be $\frac{50}{500}$, or 10 percent per annum. It is evident in this case also that the greater the price paid for the security, the smaller would be the rate of interest earned.

What Determines the Rates of Interest?

It is evident from the above that what determines the price of a given expectation also determines the rate of interest in that particular form of investment. However, the theory of Supply and Demand states that the price of any given expectation in a competitive market will be that

price at which the amount offered is equal to the amount demanded. For each class of security we can draw supply and demand curves and find the equilibrium price and quantity as in Fig. 17, Volume I. Or we could determine the equilibrium price of any security by constructing its "total market curve" (see Volume I, Chapter 7), and finding where the excess demand or supply amounted to zero. The price of any class of security is determined by the "height" of its total market curve, which again is a measure of the general eagerness to buy it. "Why is there a rate of interest?" then, is exactly the same question as "Why is the price of a security less than the total of future payments which are expected from it?" Why are not people so eager to buy securities that their price rises to the point where they bear no interest? Why, for instance, does not the price of promissory notes which promise to pay $105 next year stand at $105 now, instead of, say, $100?

The Principle of Equal Advantage Applied to Securities

To answer the above questions, we must invoke another principle of price theory,—the principle of equal advantage. If capital is mobile, the *total advantages* which result from holding it in all its various forms must tend to be equal. The enjoyment of interest on capital is not the only advantage, however, accruing to its owner. The various forms of capital differ in many ways, and these differences account for the varying rates of interest to be obtained among them. Just as differences in money wages in various occupations of labor may exist permanently, so differences in rates of interest in various occupations of capital may continue. Pleasant and convenient occupations, like teaching, which possess large nonmonetary advantages, receive relatively small money wages. Similarly, convenient and desirable occupations of capital receive small rates of interest. Generally speaking, the most desirable attributes of capital are first, safety, and second, what is called *liquidity*—i.e., easy exchangeability for other things. Suppose we had to choose among three ways of holding $1000 worth of capital. We might hold it in the form of cash, in the form of a United States government bond, or in the form of a share in a Ruritanian oil company. If, now, the rates of interest we expected to receive were equal in all these cases, it is probable that we should prefer to hold our $1000 in the form of cash, for cash possesses greater liquidity than United States bonds, and certainly greater safety than a Ruritanian oil share. The price of United States government bonds would have to be low enough to afford us a certain small rate of interest before we should be likely to prefer holding our capital in that form. The price of Ruritanian oil shares would probably have to be very low indeed before we should be tempted to hold our capital in that form.

Liquidity Preference

It is evident that the rate of interest which capital enjoys when it is held in a form not wholly liquid or perhaps not wholly safe is something in the nature of a "bribe" to induce capital to remain in that form. In the same way, the high wages of a steeple jack are a bribe to induce people to enter a dangerous and rather unpleasant occupation. If people are to hold capital in an illiquid form, i.e., in the form of securities or of real property, the price of these things must be low enough to enable the holder to obtain a certain rate of interest. If this were not so—if, for instance, the price of securities were so high that they yielded a zero rate of interest—most people would not be willing to hold their capital in this form. That is, they would try to sell their securities—to exchange them for cash, which also bears a zero rate of interest, but has advantages in the form of liquidity and safety which securities do not have. The result of this rush to sell would be, of course, a fall in the price of securities, which is the same thing as a rise in the rate of interest if the expectations which the securities represent do not change. The rise in the rate of interest would go on until, on balance, there was no desire to transfer capital holdings from the form of securities to the form of cash. At this point the price of securities would be that at which the quantity offered was equal to the quantity demanded.

A Change in Liquidity Preference

If there is a change in liquidity preference in the market—i.e., a change in the general estimate of the desirability of holding capital in the form of money—this will be reflected in a change in the price of securities and in the rates of interest which they bear. Suppose that for some reason —let us say, a war scare—the holding of money suddenly becomes more desirable. That is to say, there is an increase in liquidity preference. People as a whole wish to hold more of their capital in the form of cash, less in the form of securities. The result is greater eagerness to sell securities coupled with less eagerness to buy them. The demand curve for most securities will fall (shift to the left). The supply curve will rise (shift to the right). The result will be a fall in the price of the security. This means, however, a rise in the rate of interest. That is, when holding money becomes more desirable and holding securities less desirable, a greater bribe must be paid to induce people to hold securities instead of money.

The "Market Identity" and Liquidity Preference

This phenomenon can be expressed very conveniently in the form of the "market identity" (see Volume I, Chapter 7). If P_b is the price of a

given kind of bond (i.e., a given set of future obligations), M is the quantity of money in the market, B the quantity of these bonds in the market, r_m the proportion of total asset values which people wish to hold in the form of money, and r_b the proportion which they wish to hold in the form of bonds, then

$$P_b = \frac{Mr_b}{Br_m}$$

Other things being equal, then, the greater the quantity of money in the market, and the greater the preference for the bonds, the higher will be their price and the lower will be the rate of interest on them. The greater the quantity of the bonds on the market and the greater the preference for money, the lower will be the price of the bonds and the greater will be the rate of interest. This, in essence, is the "liquidity preference" theory of interest, which is nothing more or less than an application of price theory to the price of bonds.

Models of Liquidity Preference

The market identity also points up some of the weaknesses of the liquidity preference theory of interest. The market identity by itself is not a "model" of the bond market. We must assume at least four other behavioral equations or assumed constants in order to determine the five quantities of the market identity. In its least sophisticated form the liquidity preference theory assumes that B is given, r_b and r_m are likewise given by the psychology of the market, and that hence P_b, and, therefore, the rate of interest, is a simple function of M—i.e., increasing the quantity of money will result in a decline in the rate of interest. This is assumed in the model which underlies Keynes' *General Theory,* in which investment is assumed to be a function of the rate of interest, and the rate of interest is a function of the stock of money.[5]

Weakness of Liquidity Preference Models

The weakness of this model of the bond market is that it wholly neglects possible repercussions of an increase in the money stock on the other variables. In short-run dynamic considerations the impact of changes in price on the preference parameters, r_b and r_m, is very im-

[5] The model consists then of the following equations:
The basic identity, $P = A + C$
the consumption function, $C = F_c(P,i)$
the investment function, $A = F_a(i)$
the liquidity function, $i = F_i(M)$
The stock of money, M, being given, we have a system with four equations and four unknowns, P, A, C, and i.

portant. Thus rising prices of bonds may increase r_b and diminish r_m and so contribute to the rising price itself. Even more important, however, from the point of view of long-run equilibrium, is the possibility of changes in B—i.e., in the total volume of debt outstanding. We can postulate a "normal" equilibrium rate of interest in any particular kind of security, r_e. If the market rate of interest, as defined by the market price of the security, is above r_e we may suppose that the quantity of the security in the market will fall, since it will not be profitable to create new securities of this type, because as old securities mature and disappear from the market they will not be replaced, so that the total stock of the securities will continually decline. As this stock declines, however, the price will tend to rise, and the market rate of interest, to fall. When the market rate comes down to the normal rate it will be found profitable to create enough new securities just to replace those which are maturing, and the stock of the securities—and therefore their price—will remain constant. We are assuming here, of course, that the quantity of money and the asset preferences do not change. Similarly, if the market rate of interest on any particular kind of security is below r_e, more securities will be created than are destroyed, and the total stock will rise; as it rises, the price falls, and the rate of interest rises until once again it is just sufficient to maintain the stock constant.

The Normal Rate of Interest

It will be observed that the process described above is exactly analogous to the equilibrium of normal price of a commodity (see Volume I, Chapter 10). If the market price of a commodity is "too low," its stocks will decline, which will raise the market price; if the market price is "too high," stocks will increase, which will lower the market price. The normal price is that which adjusts production and consumption to the point where stocks are constant, or rather where they grow at a rate appropriate to the whole growth of the economy. Similarly, the normal rate of interest on any kind of security may be defined as that which will adjust the creation and destruction of these securities to the point where the rate of growth of the total stock of the security—that is, the excess of creation over destruction—is appropriate.

The Supply of Securities

The next question which arises is, What determines the net amount of securities created? as this question must be answered before we can give a satisfactory answer to the question, What determines the normal rate of interest? It is clear that the willingness of people to create and to hold debt depends in large measure on the rate of profit on real asset

combinations. If the rate of profit is high the normal rate of interest on various types of securities is also likely to be high, for if this rate is low it will pay to borrow—that is, issue and sell securities in order to obtain control of more real assets.

Thus, in the previous example we saw that if the rate of profit was 10 percent and the rate of interest was 5 percent per annum, a firm which started with assets of $1000 and liabilities of $600 would find its net worth growing at a rate of 17.5 percent. Suppose now that the firm borrowed $1000 without change in the rates of profit or interest. Assets would now start at $2000, liabilities at $1600, with net worth still remaining at $400. In a year assets would grow to $2200, liabilities to $1680, net worth to $2200 − $1680, or $580. The rate of growth of net worth is $\frac{180}{400}$, or 45 percent per annum. It is clear from the equation in footnote 4 (page 133) that the more we borrow, the greater the rate of return on our net worth as long as the rate of profit is greater than the rate of interest. Indeed, if there were no uncertainty it would be impossible for the rate of profit and the rate of interest to diverge for long; if the rate of profit under these circumstances were to exceed the rate of interest even slightly, there would be a strong incentive to borrow (i.e., to create and sell securities) and the new securities coming on to the market would depress the price of all similar securities and would raise the rate of interest until it was equal to the rate of profit. The expansion of assets might also have an effect in reducing the rate of profit, so that the divergence between the rates of profit and of interest would be eaten away from both ends.

Effect of Uncertainty

Because of the presence of uncertainty in all economic affairs, however, rates of profit and rates of interest may diverge even in equilibrium under perfectly competitive markets. The more we borrow, the greater the risk of losing our net worth if things turn against us and the rate of profit actually turns out to be negative. Thus, suppose in the above example the rate of profit actually turned out to be minus 20 percent. With initial assets at $1000 and liabilities at $600, assets would decline in a year to $800, liabilities grow to $630, so that net worth would decline from $400 to $170—a serious matter enough! If, however, the firm borrowed an additional $1000, so that it started with assets of $2000 and liabilities of $1600, assets would decline to $1600, and liabilities grow to $1680; the net worth would be − $80, that is, the firm would be insolvent. It is clear that the larger the proportion of liabilities to net worth, the greater is the chance of insolvency if things do not go well. Consequently, even

if the most likely expectation of the rate of profit is greater than the rate of interest, the businessman may still hesitate to expand his loans and his assets because of the increasing risk of insolvency, or of irrecoverable loss.[6]

Loan Rationing

Another important reason for divergence, even in equilibrium, between profit rates and interest rates is the existence of *rationing* in the loan market, especially in the market for bank loans (see page 111). Conditions which in a more competitive market would be reflected in a rise in the rate of interest (that is, conditions which necessitate a contraction in the creation of new debt) may, in fact, simply be reflected in more severe collateral[7] requirements or by a stiffening in the requirements for the granting of loans.

Normal Rate of Interest and the Rate of Profit

Even though there is likely to be a divergence, in equilibrium, between the rate of profit and the rate of interest, it is clear that if this divergence is fairly stable, then the main force determining the rate of interest in the long run is the rate of profit. This, indeed, is the core of Adam Smith's criticism of the liquidity preference theory of the mercantilists. Nevertheless, the short run may be a long time in the affairs of men, and we cannot afford to neglect it. The short-run, liquidity preference theory of the capital market and the long-run theory are not, therefore, competitors, but complement each other.

Interest and the Banking System

With the analysis developed to this point, we can now proceed to examine an extremely important aspect of the theory of the securities market, which is the impact of the banking system on the rate of interest and on the creation of securities. Commercial banks differ from other economic organizations in that their deposit liabilities constitute liquid assets, or money, for those who hold them as assets. We have seen also (page 106) that when securities are sold to banks, the total of bank deposits increases by an amount roughly equal to the increase in the bank's holdings of securities, less any losses of reserves or other assets. Consequently when a bank makes a loan, not only is a security created (the promissory note) and put on to the market, but a deposit (money)

[6] This is the principle known as the *principle of increasing risk*. See M. Kalecki, "The Principle of Increasing Risk," *Economica*, November, 1937, p. 440.

[7] Collateral consists of borrower's assets which are pledged to the lender and which become the property of the lender if the borrower fails to meet the conditions of the loan contract.

is also created. When a nonbank institution or an individual makes a loan, a security is created (the promissory note, or bond of the borrower), but no money is created directly; money is merely transferred from the lender to the borrower. An expansion of bank loans, therefore, may have an effect on the price of securities in general very different from that of the expansion of other securities. The creation of securities and their absorption into the market (i.e., the finding of people to hold them other than their original issuers) will generally depress the price of securities and so raise rates of interest. The sale of securities to banks, however, also results in the creation of money, which will in large part offset, or will more than offset, the tendency to lower the price of securities.[8] A good illustration of the above principles is found in the discussion of war finance (pages 241–244).

The Wicksellian Model

The simultaneous creation of securities and of money through the banking system leads to the possibility that a divergence between the market rate and the normal rate of interest might lead, not to a change in the stock of securities which will bring the market rate back to the normal rate again, but to simultaneous changes in the stock both of

[8] Let M be the quantity of money, B the stock of securities, r_b the preferred security ratio, and r_m the preferred liquidity ratio. The price of securities, P_b, is given by the market identity,

$$P_b = \frac{Mr_b}{Br_m}$$

Now suppose that securities increase by dB, and that as a result of these being sold to banks, the money stock (deposits) increases by dM. Suppose preference ratios do not change; then the new price of securities, P'_b, is given by

$$P'_b = \frac{(M + dM)r_b}{(B + dB)r_m}$$

It follows that
$$\frac{P'_b}{P_b} = \frac{(M + dM)B}{(B + dB)M} = \frac{1 + \dfrac{dM}{M}}{1 + \dfrac{dB}{B}}$$

We have therefore
$$P'_b \gtreqless P_b, \text{ according as } \frac{dm}{M} \gtreqless \frac{db}{B}$$

That is, if the proportionate increase in money is greater than the proportionate increase in securities, the price of securities will rise and the rate of interest fall. If now we have, as will be approximately the case where the new securities are sold to banks, $dM = dB$, the issue of the securities will cause a fall in the rate of interest if M is smaller than B.

securities and of money which will lead to general inflation or deflation. This is a view held by the great Swedish economist, Knut Wicksell.[9] If the market rate is below the natural rate, there will be net creation both of securities and of bank deposits. If now commodity prices are flexible, there will be an increase in commodity prices and, indeed, a rise in the whole wage-price level, but there may not be a movement of the market rate of interest up toward the normal rate. Indeed, the movements may even tend to *widen* the divergence between the market and the normal rate. The price inflation will raise rates of profit on real investment, and so will increase the normal rate; the increase in bank deposits may outweigh the influence of the increase in securities, so that the price of securities themselves may rise and the market rate fall. Similarly, if the market rate is above the normal rate, the result may be a shrinkage in securities and in bank deposits and general deflation without any tendency for the gap to be filled—the normal rate declining with the price deflation, and the market rate rising with a falling price of securities. It must be emphasized that these perverse dynamics are not *necessary,* in the sense that models can easily be constructed which possess much greater dynamic stability. Nevertheless, there is a great deal of historical evidence to suggest that processes like those outlined above do go on, at least for short periods. The inflationary process may be brought to an end by the banking system bumping up against legal or conventional reserve requirements and inelastic reserves, for as the expansion proceeds, of course, reserve ratios fall. The deflationary process may be more difficult to stop, especially if the price-wage level is sticky and the deflation consequently takes the form of a reduction in output and employment. Profit levels in a sharp deflation fall even below zero, and there is no sharp upper limit on the growth of reserve ratios, or of liquidity ratios in general, in the way that legal or conventional requirements impose a sharp lower limit.

SAVING AND INTEREST

We are now in a position to bring together the two parts of this chapter and to ask ourselves what are the relationships between saving, investment, output, and interest rates. Some misconceptions must first be cleared out of the way. If we define saving, in real terms, as the difference between real income (or output) and consumption, and investment as real accumulation, it is clear that there cannot be any divergence between saving and investment, as they are identically equal. The theories, therefore, which regard the rate of interest as that quantity in the system which makes

[9] Knut Wicksell, *Interest and Prices,* trans. by R. F. Kahn, London, Macmillan, 1930.

saving and investment equal must, to say the least of it, be very carefully interpreted. It does not make much sense, for instance, to draw supply and demand curves for "savings," the supply curve relating saving to the rate of interest, the demand curve relating investment to the rate of interest, and then to suppose that the equilibrium rate of interest is where these curves intersect. Such curves imply that income is independent of the volume of investment and consumption, which we know cannot be the case. Similarly we must beware of theories which conceive saving as the process of piling up liquid assets (money) which is then "invested" by being spent on investment goods. Money, as we have seen (page 83), is not accumulated by hoarding (i.e., not spending) and it cannot be got rid of by spending it! The surge of money stocks from household into business balances and out again is an important phenomenon, but it may be doubted whether it has anything much to do with the problems of interest or of investment except by its indirect effects.

The Rate of Interest Affects Consumption and Investment Functions

Nevertheless, the rate of interest—which we take as a symbol of the whole condition of affairs in the market for securities—may have substantial effects on the major components of the system, through its possible

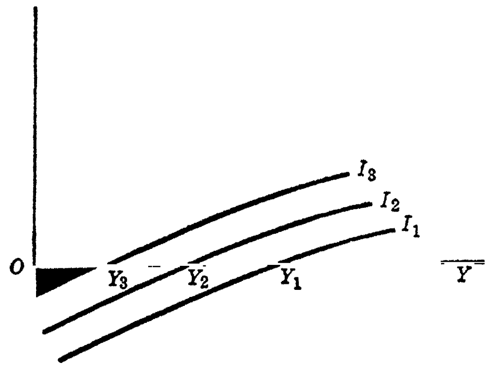


Fig. 12. Interest and Employment

effects on the consumption and investment *functions*. We may suppose, for instance, that the lower the rate of interest (i.e., the easier are the terms on which new securities can be created), the greater will be the amount of willing accumulation at each level of income—i.e., the higher the level of the investment curve.

We may if we wish also postulate an interest component of the con-

sumption function, though there is not much empirical evidence for such, except perhaps in the case of farmers and small businessmen who have opportunities for investment at high rates of return in their own enterprises. Whether we do this or not, we can combine these functions, as in Fig. 7, page 54, into an excess accumulation function, and suppose that the excess accumulation at each level of income (output) is also a function of the rate of interest. This is illustrated in Fig. 12. Here we suppose that I_1 is the excess accumulation curve at a low rate of interest; I_2, at a higher rate, I_3, at a still higher rate. We postulate intermediate curves, of course, at each possible rate of interest. Then at a rate of interest corresponding to the curve I_1, total equilibrium output is OY_1; at the higher rates output is smaller, (OY_2, OY_3). The critical question here is a factual one: whether the interest effect is large enough to make the interest rate important as an instrument of policy. If by lowering interest rates we could push equilibrium output from OY_2 to OY_1 in the figure, clearly interest and monetary policy would be very important in establishing full employment. If, however, a fall in the interest rate produces little impact on the excess accumulation curve, as seems to be the case in highly developed economies, the possibilities of interest rate manipulations seem rather meager, and we must fall back on the cruder but more powerful weapons of public finance.

QUESTIONS AND EXERCISES

1. How, in terms of the analysis of the basic model (page 49), could you interpret the suggestion that unemployment arises because saving is done by one set of people and investment by another set?
2. Suppose an economy with two sectors only, households and businesses. Suppose now that households, on the whole, exhibit a sudden desire to hold more money and to that end they cut down their purchases from businesses. Trace the immediate effects of this set of decisions: (a) on the balance sheets of households and of businesses, (b) on the demand for bank loans on the part of businesses. Trace the possible effects on output and employment on the assumptions (a) that the supply of bank loans is very inelastic (bankers are reluctant to expand their loans), and (b) that the supply of bank loans is highly elastic. Make clear the various assumptions involved. Repeat the exercise, starting from the assumption that households suddenly desire to hold less money. Repeat the exercise again, on the assumption that this time the initial impulse comes from an increased desire to hold money on the part of *businesses*.
3. What is meant by saying that the money supply is "elastic" or "inelastic"? What are some arguments in favor of, or against, an elastic supply of money?
4. "The ability of people to hoard depends entirely on the elasticity of the money supply: the ability of people to save depends entirely on the amount

of real capital which is being accumulated." Discuss the truth of this statement.

5. "Income (production) is only necessary because of consumption; the more we can economize consumption the better off we are; hence the less income we have the better off we are." Can you resolve this paradox?
6. What is the distinction between *depreciation* and *consumption?* Is this distinction (a) clear, (b) important? If it is important, for what purposes is it important?
7. Is it possible to reconcile the view that we "save in order to consume later" with the view that we save in order to build up a permanent stock of capital?
8. Under what circumstance may the increase in the total of individual net worths *not* be equal to the total value of the increase in real capital of a society?
9. What are the possible ethical implications of the proposition that the increase in the net worth of an individual is not always equivalent to an increase in the total net worth of all individuals?
10. Why do you suppose that Adam Smith and Ricardo rejected so vehemently the notion of the "paradox of thrift"?
11. Explain carefully why the age *distribution* of a population might affect (a) its actual output, (b) its capacity output.
12. "The *existence* of a rate of interest can be explained in the same way that we explain *differences* in wages." Explain, and discuss critically.
13. "The liquidity preference theory is not a theory of interest, but a general theory of prices." Discuss.
14. What similarities and differences are there between the creation and destruction of securities and the production and consumption of commodities?
15. What elements in the economic system give rise to the suggestion that a divergence between the market and the normal rate of interest gives rise to changes in the *general* price level rather than to changes in the rate of interest itself? Why cannot the same argument be applied to a divergence between the market and the normal price of commodities?
16. What is the relation of the theory of saving to the theory of interest? Is it true to say that they are related mainly through the theory of employment?
17. "The problem of the rate of interest is exactly the same problem as that of the present value of a source of income." Discuss critically.
18. During a year Mr. Smith earned $5000 salary, and received $4500 of this in cash and $500 in the form of retirement premiums. He spent $4000 on personal goods and services, and received $1000 from the sale of his automobile. He consumed $3800 worth of personal goods and services during the year, including $100 worth of garden produce which he raised himself. He purchased, in addition, $1600 worth of government bonds. Calculate the following:
    a. His money income
    b. His money outgo
    c. The total addition to his net worth during the year (saving)
    d. His total money receipts

e. His total money expenditures
f. The addition to (or subtraction from) his money assets (hoarding or dishoarding)
g. The total addition to his other (nonmoney) assets

Are there any other changes in his balance-sheet items during the year which may be deduced from the above information?

Suppose that his balance sheet at the beginning of the year consisted of the following items: cash, \$900; car, \$1000; retirement equity, \$5000; bonds, \$1000; other goods, \$5000. Construct his balance sheet at the end of the year, and show that (a) his net worth has increased by the amount saved, (b) his cash holdings have diminished by the amount dishoarded.

19. Suppose a firm in which the net worth is \$100. Suppose that the firm can borrow in unlimited amounts at 5 percent per annum. On a graph with r_n (the rate of return on net worth) measured on the vertical axis, and r_a (the rate of profit on total assets) measured on the horizontal axis, draw a series of lines showing the relationship between r_n and r_a for various values of total liabilities, L, say for $L = 50$, 100, 200, 300, 500, 1000, 2000. (Note: the r_n axis should run from about $r_n = .1$ to $r_n = -1$; the r_a axis, from $r_a = .1$ to $r_a = -0.7$.) What is the significance of this family of lines? Why do they all go through the point $r_n = .05$, $r_a = .05$? What is the significance of the points of intersection of these lines with the line $r_n = -1$? Suppose now that the firm feels that it can not possibly stand a value of r_n less than -0.3, and that it fears that the *worst* possibility during the coming year is $r_a = -0.2$. What is the maximum amount it can borrow, assuming that it cannot possibly risk an r_n of less than -0.3? Does this mean that borrowing is limited only by the fear of loss, and not by the hope of gain?

CHAPTER 7

INTERNATIONAL ECONOMIC RELATIONS

THE NATURE OF INTERNATIONAL TRADE AND INVESTMENT

The topics which are usually grouped under the heading of the "theory of international trade" are in fact merely special cases of general economic analysis. Some of these cases are treated in Volume I: the theory of comparative advantage, the theory of the foreign exchanges and the gold standard, and the "microeconomic" theory of tariffs.[1] It will be convenient, however, to group together a number of "macroeconomic" problems which arise either because of the fact that all economic activities are located at definite points on the earth's surface, or because of the existence of independent national states.

"International Trade" as a Sample

Let us examine first, then, the nature of the economic aggregates which constitute the subject matter of the macroeconomic part of the theory of international trade. By international trade we generally mean that part of the total geographical flow of goods and services which happens to cross national boundaries. It must be regarded, therefore, as a sample, and by no means a random sample, of the total volume of trade. It is a sample which acquires significance, in spite of its biased and unrepresentative nature, for two reasons. One is that the almost universal imposition of customs barriers at national frontiers leads to the collection of taxes and also of information about the movement of goods across these lines. The other is that different nations have different monetary, financial, and organizational systems and that hence the movement of

[1] Pages 27, 40, 60–65, 144–154, 211–215.

goods across international boundaries has peculiar effects which are not noticed in the case of movements within a country.

Peculiarities of International Trade

Consider, for instance, the difference between a shipment of goods from Detroit across the river to Windsor, Ontario, and a shipment from Detroit to California. The first is international trade, no matter how trivial or local in importance. The second is not international trade, although its magnitude and its consequences may be much more important. Nevertheless, the shipment from Detroit to Windsor has consequences which are not involved in the other shipment. In the first place it is recorded at the customs office, and will appear in statistical compilations. The movements of internal trade are not so recorded, and are in consequence much more difficult to trace. Then customs duties may be collected on it, with consequences for the fiscal system of the importing country. This peculiarity is not perhaps as fundamental as might be thought, for customs duties may be regarded as an increase in the cost of transport,[2] and as costs of transport have to be paid anyway, even in the case of internal trade, no new principle is introduced into the problem if we except certain difficulties which arise because the cost of transport is only changed in one direction. Thus customs duties may be thought of as simply increasing the distance between nations, and the duty itself is what has to be paid to the government to transport goods over this artificial barrier.

More important, perhaps, is the fact that the movement of goods across the boundary will have certain monetary consequences: it may leave Americans, for instance, in the possession of more Canadian dollars than they wish to hold, and so will have an effect on the rate of exchange of the two currencies. If this rate of exchange is "pegged" by a gold standard or by an exchange equalization fund, there may be secondary consequences in the form of transfers of gold or securities between different national accounts, which may have various consequences depending on the various national monetary and financial policies and institutions.

"International" Trade Problems as Distortions

The peculiarities of international trade, important as they are, should not, however, blind us to the importance of looking at the total network of trade, internal as well as international, and not merely at that sample of the total network which happens to cross the national frontiers. The problems of international trade in the narrow sense are problems of *distortion:* frontiers, tariffs, and national economies create distortions in the total network of trade, but do not create the network itself. It is

[2] See Volume I, Chapter 11.

not implied that these distortions are necessarily "bad." In thinking about this field, however, we must rid ourselves of the habit of regarding nations as homogenous "persons," called "France," "England," and so on, trading around a table! It is true, of course, that with the increased socialization of trade the picture of a nation as a trading body becomes more true to life. Even here, however, unless socialization is complete, the nation as a trading body must be regarded simply as one trader among many participating in the total volume of trade. In fact, the international trade of any particular nation is usually a collection of completely heterogeneous flows, the aggregates of which may not only lack significance but may be quite misleading. To give but a single example. Before World War II, to judge from the appearance of the aggregates, France was virtually self-sufficient in foodstuffs, exporting in normal years just about as much as it imported. The aggregate, however, covered up a great regional diversity. Most of the French import of food was from North Africa into the south of France; most of the export was from the north of France to the rest of northern Europe. Communications between the north and south of France are poor, and trade is not great. Hence, when World War II disrupted communications, the south of France starved, being cut off from North Africa, its economic partner. Normandy had food enough, being cut off from its normal export markets, but the food of Normandy could not be applied to the starvation of Provence.

Foreign Investment

Another important aggregate, or set of aggregates, encountered in international trade theory is *foreign investment.* By this is meant, in general, the increase in that part of the net worth of the inhabitants of a country which is represented by goods situated outside its borders. The *gross foreign capital* of a country may be defined as the total equity of its inhabitants in real capital (goods) located outside its borders. This is offset by the *gross foreign liabilities* of the country, this being the total equity of foreigners in the real capital located within its borders. The *net foreign capital* is the gross foreign capital minus the gross foreign liabilities. This is likely to be positive in the case of an "old" country which has been engaged in exporting capital for a while; it will be negative in the case of a country which has been importing capital. Foreign investment is the process by which the net foreign capital of the capital-exporter increases and, of the capital-importer, declines. This process takes place in many ways, both voluntary and involuntary. Thus an American corporation owned by Americans may build a factory in Brazil; a Brazilian corporation may sell shares or bonds on the New York stock market to Americans; a Brazilian importer may finance his stocks

of goods in Brazilian warehouses by discounting bills in New York. All these operations represent voluntary foreign investment by Americans in Brazil; there is an increase in the equity of Americans (stockholders, bondholders, discount houses, etc.) in goods located in Brazil. A revaluation of a currency represents considerable involuntary international redistributions of equity, just as an inflation redistributes equity internally—the owners of the depreciated money, or claims in the depreciated money, suffering for the benefit of the owners of real goods.

International Claims on "Assets in General"

Not infrequently foreign investment results, not in claims to *particular* foreign assets, but on the assets of a country in general. Intergovernment debt is usually of this character. Thus, if the United States makes a loan to Britain, this must be regarded as a vague kind of claim to British assets in general. The vagueness of these claims is reflected in their uncollectability, but as long as they are recognized, they represent redistributions of equity in the international accounts. More important, liquid assets (money) must also be regarded as a claim on assets in general of the issuing country. Thus, whoever has "dollars," whether bills or bank balances, has a claim on any American assets (goods) which are in the market, and he can exercise (and extinguish) that claim at any time by purchasing American assets with his dollars.

Trade and Investment

With this concept of foreign investment in mind, the relation between trade and investment becomes clearer. Suppose, for instance, that an American manufacturer exports automobiles to France. As long as he retains title to the automobiles on French soil, the exports are clearly also foreign investment—i.e., their value represents an increase in American equity in goods located in France. If now he sells the automobiles and receives a deposit at a French bank in francs, no change is made in the foreign investment position apart from some possible revaluations, for, as we have seen, francs represent a claim in general on French assets. If now he exchanges the francs for dollars previously held by a Frenchman, the American's claim in general on French assets is extinguished, but at the same time a Frenchman's claim in general on American assets is also extinguished, so that the international investment position remains unchanged. Trade, therefore, always represents an export of capital from the exporting country to the importing country. It follows that the gross export of capital from a country is equal to its "balance of trade"—i.e., the excess of exports over imports. An excess of imports over exports likewise represents a gross capital import. In calculating net capital import or export, allowance must be made for capital consumption and deprecia-

tion. It is not, however, the depreciation of the goods actually imported or exported which is significant, but the depreciation of goods owned by foreigners. Thus net capital export of country A is equal to A's balance of trade plus depreciation of goods in A owned by foreigners, minus depreciation of goods abroad owned by A's nationals.

The Balance of Trade

In equating the balance of trade with gross foreign investment, it must be remembered that not only imports and exports of more or less durable goods are to be counted, but imports and exports of services. Thus, when an American renders, say, a shipping service to a Frenchman, he receives, in the first instance, French money in payment, which, as we have seen, constitutes a general claim on French assets. The effect is no different, from the point of view of the volume of international capital movements, from that of the sale of a durable good. Durable goods themselves depreciate, and their value is transferred, if they are part of a profitable complex, to other assets. Services differ from durable goods only in that they depreciate more rapidly.

The Balance of Payments

The balance of trade concept must be distinguished carefully from that of the balance of *payments*. The difference is roughly similar to the distinction we have made in the case of an individual between saving (increase in total assets) and hoarding (increase in liquid assets). Thus, if a nation has a positive balance of trade, this means that its nationals are on balance increasing their total assets held abroad. If it has a positive balance of payments, it means that its nationals are increasing net *liquid* foreign assets. The situation is complicated, of course, by the existence of many forms of liquid assets—for example, gold, different national currencies, and even different forms of a single national currency. The existence of these forms makes it necessary to have a foreign exchange market, in which these various forms of liquid assets can be exchanged one for another. The theory of the determination of foreign exchange rates, is part of general price theory.[3] In so far as there are competitive markets, the rates will be determined at the level at which the market is cleared—i.e., at which the holders on balance are content to hold the amount of each currency in existence. The situation is somewhat complicated—and made potentially rather unstable—in the foreign exchange market by the fact that only a small proportion of national liquid assets at any one time (that is, those balances which the owners expect to change into other currencies) can be regarded as relevant to the foreign exchange market. In times of unstable exchange rates, speculative movements can

[3] See Volume I, Chapter 9.

often bring into the market funds from these reservoirs of domestic currency, and can so distort the exchange rates away from the values which the conditions of international trade as such might determine.

The "Payments Par" of Exchange Rates

We can define the "payments par" system of exchange rates as that system at which there will be no unwanted changes in the net liquid assets of all countries. The "payments matrix" analysis of Chapter 4 helps to clarify this problem.

Let us suppose, first, that there are only two countries, America (A) and Britain (B). Suppose Table 12 shows the payments table of this two-part system for a given week, the figures being in millions.

TABLE 12. PAYMENTS PAR IN FOREIGN EXCHANGE

	A	B	Total Expenditure
A		$28	$28
B	£10		£10
Total receipts	£10	$28	

We observe immediately that this table differs from those of Chapter 4 because the expenditures and receipts are not in the same currency. As a result of the week's transactions, Americans have paid out $28 and have received £10; Britons have paid out £10 and have received $28. Americans have increased their holdings of sterling and diminished their holdings of dollars; Britons have increased their holdings of dollars and diminished their holdings of sterling. If the status quo is to be restored, Americans must exchange their accumulated sterling for the dollars accumulated by Britons. It is evident in the example that the only exchange rate which will enable both parties to restore their previous asset structure will be $2.80 = £1, at which rate Americans will get back the $28 they have lost by giving their £10 for it, and Britons will get back the £10 they have lost by giving their $28 for it. It is shown in Appendix 1 to this chapter that no matter how many countries are involved there will always be some system of foreign exchange ratios at which the status quo, as far as money holdings are concerned, can continually be restored. These may be called the "payments par" ratios.

Equilibrating Forces in Foreign Exchanges

If the foreign exchange rates are not at the payments par levels, there will be an accumulation of foreign currency in the hands of nationals of one country or another. Thus, if in the preceding example the exchange

rate had been \$3.11 = £1, by the time Britons had exchanged all their surplus dollars (\$28) they would only have received £9 for them from the Americans; hence, Americans would be left with a £1 increase in their sterling holdings, and Britons with a £1 decrease in their sterling holdings. Similarly, if the rate had been \$2.50 = £1, there would have been a net transfer of \$3 from American to British holdings by the time all the sterling had been exchanged. Obviously, these transfers will not go on indefinitely, though they might persist for short periods if individuals were willing to increase their holdings of foreign currencies. Eventually, however, there must be a change. This may take two forms. There may be an attempt on the part of the holders of the accumulating foreign currency to buy their domestic currency in the foreign exchange market. This will depress the price of foreign and raise the price of domestic currency and so bring the exchange rates back towards the payments par. It is possible, however, that the payments themselves may be adjusted. This is the way in which purely domestic balances of payments are adjusted, as we have already seen. The people who are accumulating money will eventually stop this accumulation by increasing their expenditures; the people who are decumulating will stop the drain by decreasing expenditures. This method of attaining a payments equilibrium may be found to some extent in international trade even under a free foreign exchange market; the Americans, for instance, who are accumulating pounds may increase their purchases of British goods simply on that account.

The Gold Standard

Under a gold-standard system the method of attaining equilibrium in international trade approximates that in domestic trade, as exchange rates are fixed within narrow limits by the fact that the various national currencies are freely exchangeable for gold at fixed legal prices. Hence, under a gold-standard system, if Americans are accumulating sterling, they will simply exchange it for gold at the Bank of England and bring the gold to America to be exchanged for dollars there. Similarly, if Britons are accumulating dollars they will take them to the United States Treasury and exchange them for gold, then take the gold to England and exchange it there for sterling. This operates through the "gold point mechanism."[4]

Self-Correcting Tendencies

Under the free-gold standard of the nineteenth century the movement of gold from one country to another tended to correct the conditions

[4] See Volume I, Chapter 9.

which gave rise to it. This is particularly true in a system with no central bank, where the reserves of all the private banks are kept in the form of gold. In such a case, when gold flows into the country from abroad, the reserves of the banks will increase. Consequently, if they normally keep a constant reserve ratio, each bank will find itself in a better position to expand its loans. There will, therefore, be an expansion of credit—i.e., the banks will increase the amount of their loans and investments, and the total volume of deposits will increase. If the velocity of circulation does not fall, the new deposits will increase the money incomes of the country. This in turn may increase the volume of output or it may increase prices.

Effect on the Flow of Gold

In any case the increase in the money incomes will cause people to buy more from abroad. Imports will increase. If there is a rise in internal prices, exports may be discouraged, for the things exported will have higher prices and will not sell so readily. But the increase in imports and decrease in exports will itself act to stop the flow of gold into a country. We can see this from the previous example. If Americans, for instance, are accumulating gold it is because they are buying "too little" from abroad and selling "too much," and hence paying out less than they receive. The internal inflation, whether of prices or incomes, which the gold inflow causes will, however, make Americans more eager buyers from abroad, and America will become a good place to sell both goods and securities. Similarly, the rise in American costs and prices will hamper American exports. Americans will therefore buy more and sell less abroad, and the gap in the balance of payments which caused the gold flows will close.

The Abandonment of the Gold Standard

As an effective international system, the gold standard did not survive the First World War. A regime of fluctuating exchange rates following the war was followed by a partial return to the gold standard, led by Britain in 1925, but this did not survive the great depression. After the Second World War the International Monetary Fund was set up, the main objective of which is to provide for short-run stability in exchange rates with some provision for long-run flexibility. The case for stable exchange rates is simply that international trade becomes much more risky when the trader has to contend with fluctuating exchange rates, for then a trader must reckon on changes not only in the price at which he buys or sells but also in the rate of exchange between the currency in which he has to buy and sell and his own. If an American sells a machine in Eng-

land for £100 and the rate of exchange is £1 = \$3, that means \$300 to him, if the rate is £1 = \$2 it means only \$200 to him. A small shift in the exchange rates, therefore, may easily wipe out a trader's profit.

The Insulation of National Economies

In a somewhat narrow technical sense we can say that the inability of the Western world to restore the international gold standard was due to the development of central banking and the increasing control of central banks over their respective national banking systems. Thus, in a country in which all gold movements are centralized in the central bank, that bank can "sterilize" gold movements into or out of the country by substituting securities for gold, or gold for securities, in its asset total, thus keeping unaltered the total of its deposits (which, we recall, constitute the reserves of the member banks) in the face of quite large changes in the gold stock.

Under these circumstances a simple gold-standard system could easily result in gold flows which were not self-correcting, and which would continue until the country losing gold would be forced to abandon the standard—that is, suspend the offer to exchange gold for its currency at a fixed rate.

A much more fundamental reason for the breakdown of the old international monetary system lies in the desire of individual nations to insulate themselves from depressions originating in other countries. In order to understand this phenomenon we must now examine some macroeconomic models for an open economy—that is, one in which trade and investment take place across the boundary of the system.

MACROECONOMIC MODELS OF INTERNATIONAL TRADE

The One-Country Model

It is not difficult to extend the basic model of Chapter 3 to take account of imports and exports. It is clear that imports in the first instance add to total stocks of goods within the country, and that exports subtract from this stock. In this sense, therefore, imports have the same effect as production and exports as consumption. We can therefore extend the basic identity as follows:

$$A \equiv P + M - C - X \tag{1A}$$

where A is total accumulation, or additions to stocks, P is production, C is consumption, M is imports, and X is exports. For the purposes of

the present model it will be convenient to rearrange this identity into an exports identity:

$$X \equiv P - (C + A) + M \tag{1B}$$

We can now construct an equilibrium model if we can postulate four behavior equations. Let us suppose the following:

Consumption function, $C = F_c(P)$ (2)
Imports function, $M = F_m(P)$ (3)
Exports function, $X = F_x(P)$ (4)
"Desired accumulation" or investment function, $A = F_a(P)$ (5)

The assumptions involved are not necessarily very realistic; however, we will start from them as the simplest possible case. The first three behavior equations postulate that for every level of output (real income) there are unique levels of consumption, imports, and exports respectively, given by the consumption, imports, and exports functions. The exports function is not strictly necessary, but is put in for completeness. The fourth equation postulates that at every level of real output there is some desired level of internal accumulation of real goods, given by the investment function. To establish the dynamics of the system we must further assume that excess accumulation above the level given by the investment function leads to a fall in output, and deficit accumulation similarly leads to a rise in output.

Solution of the One-Country Model

The equilibrium is determinate, as we have five equations (1 to 5) and five unknowns, A, P, C, X, and M. The equilibrium is stable if an increase in output increases actual accumulations. The solution is illustrated in Fig. 13. As in Fig. 7 (page 54), output is measured along the horizontal axis, and the components of output are measured on the vertical axis. For the sake of simplicity in the figure we add together the consumption and investment function, so that for each level of output, home absorption, $H = C + A$, and we have a "home absorption" function, $H = F_c(P) + F_a(P)$, represented by the "home absorption curve," Hh. We draw an imports curve, Mm, representing the imports function, and an exports curve, Xx, representing the exports function. We now draw the *exports requirements curve*, $X_r x_r$, as follows. At any output, say OP, where imports are Pm, home absorption is Ph, and output is also equal to Pp (p being on the 45 degree line), we measure $mx_r = hp$. Px_r is then the exports requirement, X_r, at this level of output—that is, the amount of exports required *if* the given output is to be the equilibrium output. By the construction we see that $X_r = M + P - (C + A)$, and therefore if the

identity (1B) is to be satisfied we must have $X_r = X$; that is, the actual exports must equal the export requirements in equilibrium. The point of equilibrium then is given by the intersection of the exports curve, Xx, with the exports requirements curve, $X_r x_r$, at the point x_e in the figure. At this point, and only at this point, the excess of imports over exports, $x_e m_e$, is equal to the excess of home absorption over output, $p_e h_e$. At

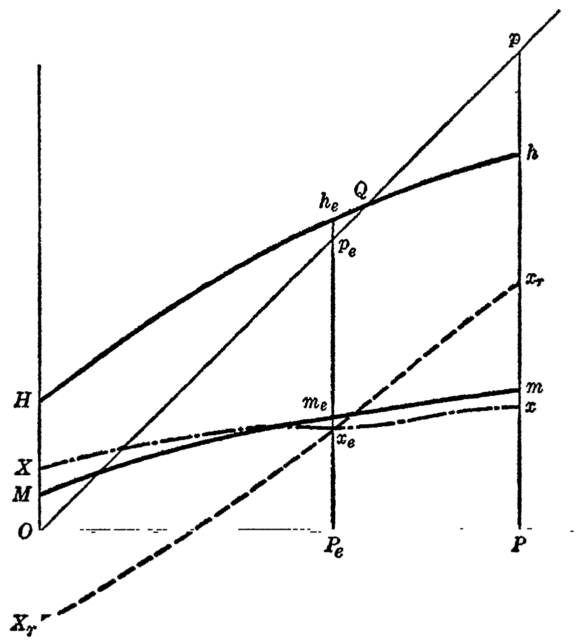


Fig. 13. Foreign Trade Equilibrium, One Country

larger outputs than OP_e there is a deficiency of exports and there will, therefore, be unwanted accumulations of goods at home leading to reductions in output. At smaller outputs than OP_e there will be an excess of exports leading to unwanted decumulations of goods at home and therefore to a rise in output.

The Export Multiplier

From this model a number of multipliers of varying degrees of significance can be derived. Suppose, for instance, that there is a spontaneous rise in the exports function, from, say, Xx_e to $X'x'_e$ in Fig. 14, leading to a shift in the equilibrium position from x_e *to* x'_e. We suppose that the

curves are straight lines within the relevant ranges. $x_e x_r$ is the export requirements curve. The *exports multiplier,* m_x, is the ratio $\frac{x_e y}{z x'_e}$ — that is, that change in the equilibrium output per unit shift in exports. Suppose α_m, α_c, α_a and α_x are the "propensities" to import, consume, accumulate, and export respectively—that is, the increase in imports, consumption, voluntary accumulation, and exports which result from a unit increase in output. We have, then,

$$zx'_e = yx'_e - yz = x_e y(1 + \alpha_m - \alpha_c - \alpha_a) - x_e y(\alpha_x)$$

that is, $$\frac{1}{m_x} = \frac{zx'_e}{x_e y} = 1 + \alpha_m - \alpha_c - \alpha_a - \alpha_x$$

or $$m_x = \frac{1}{1 + \alpha_m - \alpha_c - \alpha_a - \alpha_x} \tag{6}$$

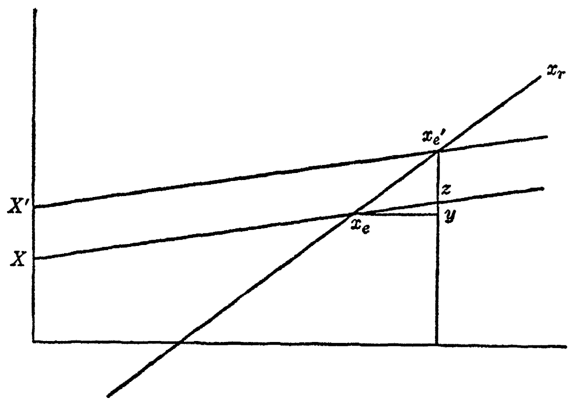


Fig. 14. The Export Multiplier

We can of course simplify the model somewhat by supposing that exports are an "exogenous" factor given by the outside world, in which case $\alpha_x = 0$, but it does not seem unreasonable to suppose that α_x in fact might be positive.

The Foreign Investment Multiplier

In a similar way we could derive multipliers for any of the other variables of the system, or even for combinations of variables. We might, for instance, define a *foreign investment multiplier* as the increase in equilibrium output which would result from a unit increase in foreign investment *(X — M).* Indeed, of the multiplication of such multipliers there seems to be no end. Whether any of these multipliers are useful instru-

ments of analysis, however, is another question, and the student must bear clearly in mind the sharp, and often unrealistic, limitations which must be placed on the system if a multiplier is to be meaningful. A multiplier describes the effect on equilibrium output of a "vertical" shift in one of the behavior functions of the system, *on the assumption that the other functions do not change.* It is the latter assumption which is so questionable in many cases, especially in the case of the foreign trade multipliers. It is most unlikely, for instance, that the imports function could shift spontaneously without any shifts in the other behavior functions. Probably the exports multiplier makes the most sense, as we might assume a shift in the foreign demand for a country's exports which did not depend much on its domestic behavior functions or have much effect on them.

The principal weakness of the above model is that it is confined to a single country, and hence is not really general enough, for the system of trade is a world system and hence the equilibrium of one country must be affected by decisions made in others. It is particularly unrealistic to assume an exports function which shows exports only as a function of income (or other variables) in the exporting country. It is necessary, therefore, to expand the model at least to include two countries. As we expand the model, unfortunately, it becomes much more difficult to analyze. These difficulties however are inherent in the subject and must be faced.

The Two-Country Model

Let us consider, then, the theory of a "two-part economy"—i.e., a closed "world" economy which is divided into two "countries." It should be observed that this model has an importance far beyond the theory of international trade: it can be applied, for instance, to the study of the relations between any two "halves" of the world economy, such as, for instance, the agricultural sector and the nonagricultural sector, or the consumption goods sector and the investment goods sector. The simplest possible model we can construct of such a two-part economy consists of six equations and six unknowns. Let P_1, P_2 be the total outputs (income) of the two sectors. Let H_1, H_2 be those amounts of product which are willingly absorbed at home in each sector. H_1, that is to say, is equal to the consumption of sector I, plus the willing accumulation (home investment) in sector I. Then let M_1 be the imports into sector I, which is the same as the exports from sector II (X_2), and let M_2 be the imports into sector II, which is the same as the exports from sector I (X_1). As there are only two sectors, anything which is an import of one sector must be an export of the other. We now postulate six equations to determine these six unknowns:

Two basic identities, like identity (1):

$$P_1 = H_1 + X_1 - M_1 = H_1 + M_2 - M_1 \tag{7}$$
$$P_2 = H_2 + X_2 - M_2 = H_2 + M_1 - M_2 \tag{8}$$

Two behavior functions for country I:

$$\text{Home absorption function, } H_1 = F_{h1}(P_1) \tag{9}$$
$$\text{Imports function, } M_1 = F_{m1}(P_1) \tag{10}$$

Two similar behavior functions for country II:

$$H_2 = F_{h2}(P_2) \tag{11}$$
$$M_2 = F_{m2}(P_2) \tag{12}$$

An algebraic solution for this system, for the special case of linear functions, is given in appendix 2 to this chapter. A graphic solution is shown in Fig. 15. Linear functions are also used here for the sake of simplicity in drafting, but the graphical method is also applicable to nonlinear functions. Output of country I (P_1) is measured along OY_1, of country II (P_2) along OY_2. The components of P_1 are measured along

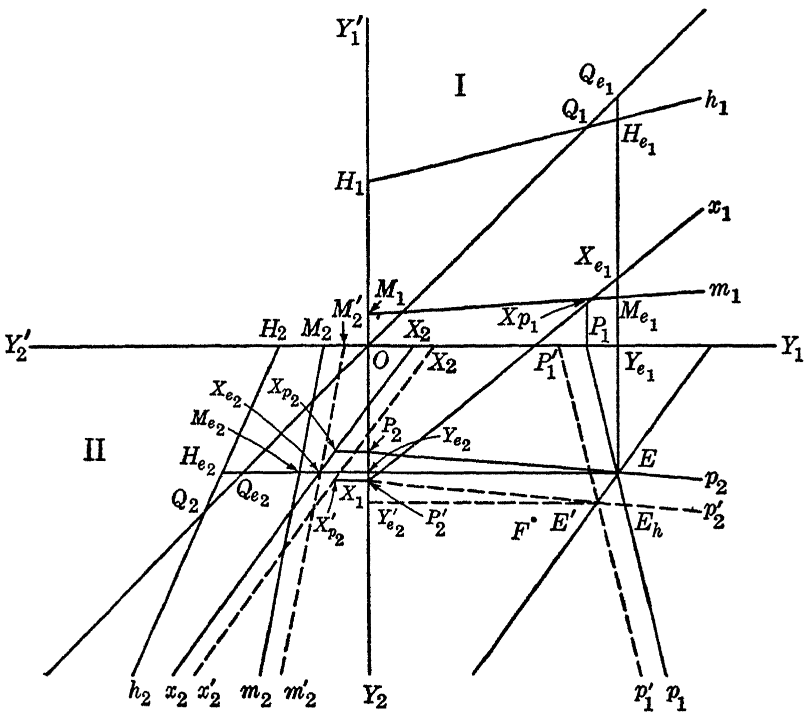


Fig. 15. A Two-Country Model

OY'_1, the components of P_2 along OY'_2. Thus the diagram in the field bounded by OY_1 and OY'_1 is the "basic model" diagram for country I, analogous to Fig. 13, page 157. The field bounded by OY_2 and OY'_2 is the similar diagram for country II; it may look a little unfamiliar to the reader because of the location of the axes, but if he will rotate the page 90° and look at it in a mirror, he will see that it is essentially the same figure as for country I. H_1h_1, then, is the home absorption (consumption plus investment) curve for country I, H_2h_2, the corresponding curve for country II. M_1m_1 and M_2m_2 are the imports curves for the two countries. We now construct the exports requirements curves, x_1X_1 and x_2X_2, for each country. In the present figure, with H_1h_1 and M_1m_1 straight lines, X_1x_1 will also be a straight line, so that it is only necessary to find two points on it and join them with a straight line. It will be observed that the export requirements curve intersects the imports curve at a point immediately below the point Q_1 where the home absorption curve intersects the 45 degree line. The distance $M_{e1}X_{e1}$ ($= H_{e1}Q_{e1}$) represents the amount of foreign investment, or the export surplus which is necessary to prevent unwanted home accumulations at the level of output OY_{e1}. That is to say, at any level of output OY_{e1}, $Y_{e1}H_{e1}$ is all that will be disposed of at home either in consumption or in willing accumulation; the excess of income over home absorption, therefore, $H_{e1}Q_{e1}$, must be disposed of as net exports if there are not to be unwanted accumulations at home forcing a contraction of output.

For country II, at the output OY_{e2} it will be observed that there is an excess of home absorption over output amounting to $H_{e2}Q_{e2}$. This must be taken care of by an excess of imports over exports if stockpiles are not to be depleted.

Equilibrium of the Two-Country Model

We now construct a "P_1 curve," P_1p_1, which shows the equilibrium position of the output of country I for *each* output of country II. This is constructed as follows. For any output of country II, say OY_{e2}, the imports of country II are $Y_{e2}M_{e2}$. This is equal to the exports of country I. To find the equilibrium output of country I, then, which corresponds to the output OY_{e2} of country II, we follow the export requirements curve of country I, X_1x_1, to the point X_{e1}, where, at the output OY_{e1}, the export requirements of country I, $Y_{e1}X_{e1}$, are equal to the imports of country II, $Y_{e2}M_{e2}$. The perpendiculars to the axes at Y_{e2} and Y_{e1} intersect at E, which is therefore on the P_1 curve. In the present case the curve is a straight line, and it is only necessary to find two points on it and join them. Thus P_1 is the point on OY_1, where $P_1X_{p1} = OM_2$.

In an exactly similar way we construct the "P_2 curve," P_2p_2; thus the

point P_2 is where the export requirements of country II, P_2X_{p2}, are equal to the imports of I, OM_1. The point where the P_1 curve and the P_2 curve intersect, E, is the point of equilibrium of the whole system, where all six equations (7) to (12) are satisfied. If the curves are linear it is clearly the only point in the field where all the equations are satisfied. OY_{e1} is the equilibrium output of country I, and OY_{e2}, the equilibrium output of II. At these outputs only are the export requirements of each country equal to the imports of the other.

Shifts in the Two-Country Model

We can now study the effect of changes in the functions of the model on the position of equilibrium. Let us suppose first that there is a "downward" shift in the imports function of country II, reflected in a shift in the imports curve M_2m_2 to a position $M'_2m'_2$, at which there is a smaller volume of imports at each level of output. The export requirements curve of country II likewise shifts downward to the position X_2x_2. The fall in the export requirements curve of country II then "raises" the P_2 curve to the position $P'_2p'_2$. At each level of output of country I there is still the same volume of I's imports and therefore of II's exports, so output in II will now be larger than before at each level of the output of I. Thus P_2 will move to P'_2, where $P'_2X'_{p2}$ is now equal to OM_1. Similarly the fall in the imports curve of II "lowers" the P_1 curve from P_1p_1 to $P'_1p'_1$, as corresponding to each output of II there is now a smaller export from I and therefore a smaller output from I. The position of equilibrium, E, therefore, moves to a new position E', at which the output of II is larger and the output of I is smaller. As long as the home absorption functions do not change, the position of equilibrium must lie on the line through EE' (in the present figure straight line). This line shows all combinations of outputs for which the excess of home absorption over output of one country (e.g., $H_{e1}Q_{e1}$) is equal to the corresponding deficiency in the other (e.g., $H_{e2}Q_{e2}$). The equilibrium point must be somewhere on this line, as the difference between the exports of one country and the imports of the other must be equal to the excess or deficiency of production over home absorption in each country—otherwise there will be unwanted accumulations somewhere in the system.

The Economic Warfare Line

The line EE' might well be called the *economic warfare line.* It shows how one nation can improve its own output at the expense of another through manipulating the import functions. Thus we saw how country II, by depressing its import function (say by imposing tariffs, quotas, or other restrictions on imports), might shift the point of mutual equilib-

rium from E to E', with the expansion of its own output and the diminution of the other nation's output. Similarly, by depressing its own imports curve, country I might move the point of equilibrium back up the economic warfare line toward E or even beyond. The reader may wonder at this point what has become of all the much advertised advantages of international trade and specialization noted in earlier chapters. In the present model it looks as if each nation has an advantage in reducing its imports, provided the other nation does not. If each nation, therefore, acts without regard to the possible effects of its actions on the policy of the other nation, it would seem as if each nation will lower its imports curve until in the final equilibrium each imports curve coincides with the corresponding output axis and there is no international trade whatever!

Retaliation

Two things modify this dismal conclusion. The first is the possibility that each nation may take into account the possibility of retaliation. Thus if country II was fairly sure that in the event of its restricting imports in the attempt to move from E to E', country I would retaliate with a sufficient restriction of its imports to bring the equilibrium back to E, there would be some hesitancy at least in starting a restrictive policy; for if both nations restrict imports the final equilibrium as far as employment and total output are concerned may be just the same as it was before, but both nations are actually worse off because the *composition* of their consumption and absorption patterns is worse. Both nations, that is to say, will lose some of the advantages of specialization and trade, and neither will gain anything in employment and output. Unfortunately, however, the fear of retaliation seems to be a very weak motivation in situations of this kind, as the rather similar case of the arms race in international political relations also indicates. The motivation is still weaker as we move from the two-nation case to the many-nation case, for whereas the possibility of retaliation in the two-nation case is quite large, where there are many nations a small nation among them may get away with import restrictions which do not provoke retaliations because their effects are so widespread.

Full Employment Limits Economic Warfare

Fortunately, however, there is a second limitation on the above model which offers some hope for the persistence of international trade. The economic warfare line is only significant within the limits marked out by the full employment or capacity outputs of the two countries. Thus, suppose in Fig. 15 that OY_{e1} represented the capacity output of country

I and OY'_{e2} represented the capacity output of country II. Then only that portion of the economic warfare line between E and E' would be significant, as the system could not move to positions beyond the capacity of either country. That is to say, when any country reached capacity it would lose any further incentive to reduce imports. Unfortunately, this does not mean that the process of competitive import restriction might not set in, even when there is only a narrow range of possibilities on the economic warfare curve. Thus country II might want to push the position of equilibrium from E to E', and country I then might push it back to E, and country II push it back again to E', all the time through successive lowerings of the import functions. If, however, the range is narrow enough, and still more if the point of full employment for both countries lies *on* the economic warfare curve, the position may be stable, especially if each country takes some account of the possibilities of retaliation, and also perhaps balances the loss from the diminution of trade as such against possible gains in employment.

Effects of Raising Home Absorption

Movement along the economic warfare curve through import restriction is not, of course, the only means by which a country can seek to expand its output up to capacity. It can also seek to raise its home absorption curve, either through fiscal or monetary policy or other methods of stimulating consumption and domestic investment. Such a course is twice blessed—if successful, it raises both the income of the nation which undertakes the policy and the income of the other country as well. Thus, suppose that country II is at a position of less than capacity output, and seeks to raise this output by measures which will "raise" the curve H_2h_2 to give a larger amount of home absorption at each level of output, without change in the imports function. The export requirements curve will again "fall," as home absorption is a negative component of the exports requirement ($P_2 - H_2 + M_2$). This will "raise" the P_2 curve, say from the position P_2p_2 to $P'_2p'_2$. There will not, however, be any change in the P_1 curve, as there is no change in the imports curve, M_2m_2, or in any of the curves of country I. The position of equilibrium then moves from E to E_h. There is an increase in the equilibrium outputs of both countries. We can, if we like, regard this as a *division* of the benefits of the policy, the international trade situation resulting in a certain diversion of the benefits of the policy from the nation which undertakes it to the nation with which it trades. The amount of this diversion depends on the slope (in this case) of the P_1 curve. Movement along the economic warfare curve is sometimes called the "export of unemployment." Similarly the diversion of the effects of a domestic employment policy through trade might be called the "export of employment."

APPLICATIONS OF THE TWO-COUNTRY MODEL

Mercantilism

The above model throws a great deal of light on the persistence of mercantilist ideas, especially in the world of "practical" men, in spite of the long and trenchant attack by the economists. The virtues of free trade are realized only in a full-employment world—i.e., a world in which the point of capacity outputs for all nations lies on the economic warfare curve, and hence there is no relevant section of the economic warfare curve at all. In a world in which the home absorption functions are not "high" enough to yield this desirable result, the economic conflict curve may have a significant range and the free-trade position, therefore, becomes theoretically, as well as practically, unstable. The political instability of a free-trade system rests also on certain other grounds. Any tariff or other restriction of imports is likely to benefit some small group, whereas the injuries are likely to be spread over the society at large. Consequently the benefits of trade restrictions are likely to have a much higher political visibility than the injuries. There is a general presumption because of this that trade restriction will be carried far beyond any possible theoretical optimum, as the political influence of vocal and organized minorities who are aware of possible benefits nearly always outweighs the influence of diffuse and unorganized majorities.[5]

Foreign Investment in the Two-Country Model

We have still not exhausted the two-country model, simple as it is. At the position of equilibrium, E, there is "foreign investment" by country I in country II. Country I has an excess of exports over imports amounting to $M_{e1}X_{e1}$, and country II has an exactly equal excess of imports over exports amounting to $M_{e2}X_{e2}$. This excess of exports of country I (or of imports of country II) is the foreign investment. It is necessary in order to take care of the excess of production over home absorption in country I, $H_{e1}Q_{e1}$, and the corresponding deficiency in country II, $H_{e2}Q_{e2}$. Now, however, the question arises, Suppose that investors are not willing to allow foreign investment to go on at this rate; what then? Suppose, that is to say, we add another behavior function to the system, to represent "willing" foreign investment. This could have many forms. If B_{12} is the net foreign investment of I in II, we might write

$$B_{12} = F_b(P_1,P_2) \qquad (13)$$

[5] There are certain grounds for believing that even in a full-employment system there may be an advantage in trade restriction for the nation imposing that restriction, much as a monopolist can obtain gains for himself at the expense of other members of society by restricting his output.

Or we might suppose simply that B_{12} is given by factors outside the system. In either case, we now have seven equations (7–13) and only six unknowns, for B_{12} is given by the identity

$$B_{12} = M_2 - M_1 \tag{14}$$

and is not, therefore, an independent unknown. The system, that is to say, is overdetermined. It would only be by accident that the value of B_{12} at the point E corresponded to the value given by the behavior equation (13). We must face the question as to what will happen to the system if there is a discrepancy between the amount of foreign investment which is necessary for equilibrium (say $M_{e1}X_{e1}$) and the desired amount at these levels of output as given by the behavior equation. If there is such a discrepancy, it is clear that some machinery must exist whereby the various functions themselves can be adjusted until the discrepancy is eliminated, and the "internal" equilibrium system, as given by the point E, is consistent with the "external" equilibrium, as given by the desired rate of foreign investment.

The Price System in the Two-Country Model

Here we see the necessity for a certain reconciliation between the "Keynesian" models of international trade which we have been discussing and the "classical" models in which international adjustments were assumed to be made mainly through the price system. If the overdeterminacy of the model noted above is to be avoided, an additional unknown must be added to the system. The most obvious addition to make is some variable which is representative of the relative price structure of the two countries. The simplest case, theoretically, is that in which the countries possess different currencies which are exchanged against each other in a free foreign exchange market. The foreign exchange rate then can be introduced as a seventh variable which can be included in the imports functions and also in the foreign investment function. Thus, suppose that there is net foreign investment of country I in country II, as at the point of equilibrium, E, in Fig. 15, and suppose that the amount of foreign investment is "too great" at these levels of output, in the sense that nationals of I are piling up claims on the assets of II in excess of the rate at which they wish to accumulate them. The nationals of I will respond to this situation by trying to exchange these claims for claims on their home assets. That is, in the market as a whole people will be trying to get rid of currency II and acquire currency I. This will move the exchange rate, making currency I dearer in terms of currency II. This means that country I will find its imports are cheaper in terms of its own currency, and its imports function will therefore rise—that is, it will take

more of the now cheaper imports at each level of output. Similarly, country II will find its imports dearer, and its imports function will fall. The point of equilibrium, therefore, will move from E along the economic warfare curve toward E'. As it moves, however, the foreign investment of I in II declines. When it has declined to the point where the actual amount of foreign investment is equal to the desired amount, the equilibrium of the seven-equation system has been reached.

Adjustments with Pegged Exchange Rates

Now let us suppose that the exchange rates between the currencies of the two countries are fixed, either because both countries conform to a gold standard, or some other pegging arrangement, or because the two parts of the economy are governed by a single monetary system, as would be the case if we were discussing the relations between two parts of a national economy. In this case the adjustments have to be made mainly by shifts in the relative price structure of the two countries or parts. In the case of the traditional gold standard, "unwanted" foreign investment will result in trade moving to the gold points, and gold will flow from the country in which the investment is made to the investing country. If now the result of the gold flow is a monetary expansion in the gold-importing country and a monetary contraction in the gold-exporting country, equilibrating forces will be brought into play. The dynamics of the situation, however, are extremely complicated, as we cannot assume that the domestic absorption functions will remain unchanged. In the case of equilibrium brought about by changes in the exchange rates, it is not unreasonable to assume, at least as a first approximation, that the equilibrating movement takes place through changes in the imports functions alone. There is in that case a movement along the economic warfare curve, as we have seen, but no shift in the position of that curve. Where, however, the exchange rates are fixed, so that the equilibrating factor has to be the general monetary changes in the two areas, changes in the home absorption function cannot be ruled out. The monetary expansion in one area may raise and the contraction in the other area may lower the corresponding home absorption function. If the rise in one is exactly similar to the fall in the other, the position of the economic warfare line will not change. The reader may confirm this by lowering the line H_2h_2 in Fig. 15 by exactly as much as he raises the line H_1h_1, and working out the position of the economic warfare line EE'. What will happen, however, is that the amount of foreign investment corresponding to any particular point on the economic warfare line changes. In the present case the point of zero foreign investment moves "up" the line toward the OY_1 axis. The amount of foreign investment

which is required at the point of equilibrium, *E*, to secure domestic equilibrium, therefore, falls as H_1h_1 rises and as H_2h_2 falls. If this process goes on far enough, equilibrium could be achieved without any change in the import functions. The situation is more complicated, of course, if the change in one home absorption curve does not exactly counteract the change in the other, so that there is a movement in the economic warfare curve itself.

Other Equilibrating Changes

It should be observed also that there are other possible equilibrating mechanisms besides the monetary changes. There may, for instance, be flows of population between the two areas, which will also have the effect of shifting the home absorption curve—raising it in the country receiving the immigrants and lowering it in the country losing emigrants. The effects here, however, are difficult to trace, and may even move the system away from equilibrium. Thus, if with the domestic equilibrium at *E* in Fig. 15 and an excess of investment of I in II, population moves from I to II—which is by no means unlikely—the movement of population may intensify the disequilibrium, raising H_2h_2 and lowering H_1h_1 and moving the equilibrium point still farther from *E*. The situation is also complicated because in practice movements of population affect the willingness to undertake foreign investment in the form of immigrants' remittances to their place of origin.

SOME SPECIAL PROBLEMS OF THE BALANCE OF PAYMENTS

Problems involving the balance of payments have attracted a good deal of attention since the end of World War II. In the years immediately after the war a "dollar shortage" was experienced, in the sense that there seemed to be a tendency for the United States export surplus to exceed chronically what would be demanded by voluntary foreign investment. This led to import and exchange restrictions on the part especially of the major European countries, whose economies had been devastated by the war, and also to persistent pressure to enlarge American government payments abroad. Some economists held the view that the dollar shortage would be a permanent disequilibrium. However, in the 1950s and 60s the dollar shortage turned into a "dollar surplus," that is, a persistently unfavorable balance of payments for the United States, with a great rise in dollar balances held abroad, especially in Europe (the so-called "Eurodollar"), and also a substantial drain of gold from the United States to Europe. In part this has been a natural "surge" phenomenon in

the balance of payments and a quite proper redistribution of the stock of international liquid assets. The movement went so far in the early 1960s, however, that the United States was forced into some mildly restrictive policies to check the unfavorable balance, and many writers expressed the fear that the total volume of international liquid assets was insufficient to prevent a long-run deflation. The situation was complicated by the fact that the dollar itself had become an important, perhaps the most important, international liquid asset, and that its status as such, depending on the willingness of foreigners to hold dollar balances as liquid assets, was in turn dependent in a rather ill-defined way on the dollar's partial convertibility into gold at a fixed price of $35 per ounce. For this reason the most obvious remedy for the persistent unfavorable balance of payments of the United States—a rise in the gold price—was not readily available, as any such change might sharply diminish the willingness of non-Americans to hold dollars and so cause a further drain on the United States gold reserve. The dilemma here is that if the structure of foreign exchange rates is not at the payments par, the dynamic effects—for instance, on asset preferences—of an attempt to change the exchange rates may be very disturbing; on the other hand, the "classical" method of adjusting balances of payments under fixed exchange rates may be ruled out by the fact that a deflation of internal prices and money incomes which might be required in some nations is virtually impossible to achieve in these days without an unacceptable volume of unemployment. It is possible, therefore, that if the equilibrium, *E*, in Fig. 15 is at less than capacity output for one or the other country, the attempt to expand output may produce difficult balance-of-payments problems.

Hyperemployment and Inflation

At the other end of the scale the various home absorption and imports functions may be such that the nations reach full employment or capacity output *before* reaching the economic conflict curve. Suppose, for instance, in Fig. 15 that the capacity outputs of both countries is given by a point, *F*, to the left of the line, *EE'*. At this point country II has a sharp excess of imports and country I, a sharp excess of exports; the movement which might bring these to equilibrium, however, would involve an expansion of output in both countries, which is by definition impossible. This is the situation of hyperemployment, and it will inevitably result in inflation. If it results in inflation in both countries, the home absorption curves in both countries should fall, which will shift the economic conflict curve toward the origin; if the shift is large enough it will go through the point, *F*, and equilibrium will be established again. As we have seen earlier, however, the effect of inflation on the home absorption function

is at least dubious, for investment may be stimulated as much as consumption is restricted; hence inflation may not be adequate to restore the equilibrium. If the inflation is suppressed by price controls the forces tending to restore equilibrium are weakened further. It is not suggested here that the above describes any actual situation. It is, however, a possible interpretation. In all these applications the economic models will be most useful if their limitations are constantly carried in mind. They are intended to be a rough sketch map of the problems of the real world, not a detailed account.

QUESTIONS AND EXERCISES

1. Suppose that a large country were broken up into a number of smaller countries. What would be likely to happen (a) to the total volume of international trade, and (b) to the total volume of all trade?
2. Enumerate the similarities and differences between (a) a shipment of goods from Buffalo to Toronto, and (b) a shipment of identical goods from Buffalo to Cleveland, in regard to as many aspects of these two transactions as you can enumerate.
3. What is meant by the "aggregation problem" in economics, and what phases of it are manifested in the theory of international trade?
4. Suppose the government of Brazil, which has previously bought and sold cruzeiros at a rate of 1200 to the dollar, devalues its currency and now buys and sells them at a rate of 1500 to the dollar. Trace as many effects of this change as you can on the *distribution* of wealth.
5. What may be meant by the "finance" of international capital movements? In what ways may these movements be financed?
6. Would it be possible for a nation to have a positive balance of trade and a negative balance of payments in the same period? Explain carefully.
7. Why is it generally regarded as undesirable to allow free markets in the foreign exchanges, but desirable to allow free markets in commodities?
8. What weakness in the gold standard system led to its abandonment?
9. Under a universal gold standard system, what effects, both immediate and ultimate, would you expect the following events to have upon the transfer of gold to and from the United States, other things always remaining the same?
   a. A great international conference in Washington
   b. A sudden increase in the number of American tourists to Europe
   c. A great boom on the New York stock market, attracting funds from abroad
   d. A sudden increase in the rates of the United States tariff
   e. A great American loan to South America
10. What would be the effects, both immediate and ultimate, of the above events on the foreign exchange rates under free exchanges?
11. In the system given by equations (1) to (5), what would be the formula for the domestic invesment multiplier?

12. What would happen to the model of equations (1A) to (5) if the capacity of the system were *below* the point of equilibrium given by the model?
13. What will be the effects on the equilibrium outputs, exports, imports, and foreign investment of the two countries in Fig. 15, page 160, of the following changes (each taken separately)? (a) an upward shift of the home absorption curve of country I; (b) an upward shift in the home absorption curves of both countries; (c) an upward shift in the home absorption curve of country I, and a downward shift in the home absorption curve of country II; (d) an upward shift in the imports curve of country I; (e) a downward shift in the imports curve of country II; (f) a combination of (d) and (e).
14. Can you construct multipliers in the case of the model in Fig. 15?

APPENDIX 1

THE PAYMENTS PAR OF FOREIGN EXCHANGE RATES

It can be shown that no matter how many separate countries there are, some system of exchange rates of their currencies can be found at which the whole system of payments will be in balance. Let the countries be A, B, C, . . . N, the names of their currencies "Alpha," "Beta," etc. Then we can construct a payments matrix between them like that in Table 13. Here, a_b represents the payments from A to B, b_a from B to A, and so on.

TABLE 13. THE PAYMENTS PAR

	A	B	C · · ·	N
A		a_b	a_c	a_n
B	b_a		b_c	b_n
C	c_a	c_b		c_n
·	·	·	·	·
N	n_a	n_b	n_c	

Let e_a, e_b, . . . e_n be the total expenditures of the various countries. Then

$e_a = a_b + a_c + \cdots + a_n$ is a sum of Alphas, $e_b = b_a + b_c + \cdots + b_n$ is a sum of Betas, and so on. The receipts of each country consist of a heterogeneous aggregate of all other currencies. The problem is what system of exchange rates will make the receipts of each country equal in value to its expenditures. Now let $\beta, \gamma \cdots \nu$ be the exchange rates of each currency in terms of Alphas; i.e., 1 Beta is equal to β Alphas, and so on. The exchange rate of Alphas, of course, is equal to 1. Then b_a Betas is worth $b_a\beta$ Alphas, c_a Gammas are worth $c_a\gamma$ Alphas, and so on. We have then, the following equations if these exchange rates are to be a payments par, representing the fact that the value of the receipts of each country, reduced to Alphas, is equal to the value of its expenditures reduced to Alphas:

$$
\begin{array}{lllll}
\text{for A:} & b_a\beta + c_a\gamma + & + n_a v & = a_b + a_c + & + a_n \\
\text{for B:} & a_b + c_b\gamma + & + n_b v & = (b_a + b_c + & + b_n)\beta \\
\text{for C:} & a_c + b_c\beta + & + n_c v & = (c_a + c_b + & + c_n)\gamma \\
 & & \cdots & & \\
\text{for N:} & a_n + b_n\beta + + & & = (n_a + n_b + & + n_n)v
\end{array}
$$

We have apparently N equations and only $N - 1$ unknowns, as α is given. However, because of the properties of the payments matrix, one of these equations can always be derived from all the others, so that there are in reality only $N - 1$ independent equations. These can be solved to discover the par values of the exchange rates, $\beta, \gamma, \cdots \nu$.

APPENDIX 2

ALGEBRAIC SOLUTION OF THE TWO-SECTOR MODEL

If the functions of the two-sector model illustrated in Fig. 15, page 160, are assumed to be linear, it is not difficult to obtain algebraic solutions

for the system. Thus suppose we replace the equations (7) to (12), page 160, by the following linear system, using the same notation: Two identities:

$$P_1 = H_1 + M_2 - M_1 \tag{1}$$
$$P_2 = H_2 + M_1 - M_2 \tag{2}$$

Two home absorption functions:

$$H_1 = H_{10} + h_1 P_1 \tag{3}$$
$$H_2 = H_{20} + h_2 P_2 \tag{4}$$

Two imports functions:

$$M_1 = M_{10} + m_1 P_1 \tag{5}$$
$$M_2 = M_{20} + m_2 P_2 \tag{6}$$

Solving these equations for P_1 and P_2 we get

$$P_1 = \frac{(H_{10} + M_{20} - M_{10})(1 - h_2) + (H_{10} + H_{20})m_2}{(1 - h_1)(1 - h_2) + (1 - h_1)m_2 + (1 - h_2)m_1} \tag{7A}$$

$$P_2 = \frac{(H_{20} + M_{10} - M_{20})(1 - h_1) + (H_{10} + H_{20})m_1}{(1 - h_1)(1 - h_2) + (1 - h_1)m_2 + (1 - h_2)m_1} \tag{7B}$$

It should be observed that H_{10}, H_{20}, M_{10}, M_{20} are the amounts of home absorption and imports in the two countries at zero outputs. That is, $H_{10} = OH_1$ in Fig. 15, $H_{20} = OH_2$, $M_{10} = OM_1$, $M_{20} = OM_2$. The coefficients h_1, h_2, m_1, m_2 are the "slopes" or the "marginal propensities" of the various functions. It is clear from equations (7) that if the propensities are constant a *rise* in the home absorption curve of either country (i.e., an increase in H_{10} or H_{20}), a *rise* in the imports curve (i.e., in M_{10}) of country II, or a *fall* in the imports curve of country I, will raise P_1, and that a symmetrical proposition holds for P_2.

The economic warfare line is given by the condition

$$P_1 - H_1 = -(P_2 - H_2) \tag{8A}$$

(That is, the excess of home absorption over output in one country must be equal to the excess of output over home absorption in the other.) Substituting the values of H_1 and H_2 from equations (3) and (4) we have

$$P_1 - H_{10} - h_1 P_1 = -P_2 + H_{20} + h_2 P_2,$$
$$P_1(1 - h_1) + P_2(1 - h_2) = H_{10} + H_{20} \tag{8B}$$

If the values of P_1 and P_2 from equations (7) are substituted in this equation it will be seen that the point of equilibrium must always lie on the economic warfare line.

Multiplier formulas can be derived, if desired, for equations (7A) and (7B). Thus we have as the "home absorption multiplier" for country I

$$m_{h1} = \frac{dP_1}{dH_{10}} = \frac{(1 - h_2) + m_2}{(1 - h_1)(1 - h_2) + (1 - h_1)m_2 + (1 - h_2)m_1} \tag{9}$$

In the absence of trade we would of course have $h_2 = m_2 = m_1 = 0$, and the multiplier would reduce to the usual formula (cf. page 57), $\frac{1}{1 - h_1}$

It is not difficult to show that foreign trade reduces the multiplier, as some of the effects of domestic expansion leak out to raise outputs abroad. Thus we have

$$\frac{(1 - h_2) + m_2}{(1 - h_1)(1 - h_2) + (1 - h_1)m_2 + (1 - h_2)m_1} \gtreqless \frac{1}{(1 - h_1)} \tag{10}$$

according as

$$(1-h_1)(1-h_2)+m_2(1-h_1) \gtreqless (1-h_1)(1-h_2)+(1-h_1)m_2+(1-h_2)m_1$$

i.e., according as $\qquad (1 - h_2)m_2 \lesseqgtr 0.$

In fact, as $(1 - h_2)$ must be positive if the system is to be stable, as m_2 is also likely to be positive, $(1 - h_2)m_2 > 0$ and, therefore, the multiplier with foreign trade [the left-hand side of inequality (10)] is less than the multiplier without foreign trade (the right-hand side of the inequality).

QUESTIONS AND EXERCISES

1. The following is a payments table for four countries:

	A	B	C	D
A		80	40	100
B	30		60	40
C	40	30		20
D	8	18	22	

Assume that expenditures are made in the national currency. Show that the payments par values of the currencies of B, C, D, relative to that of A, are 1 of B = 2 of A, 1 of C = 3 of A, 1 of D = 5 of A. Show that if all foreign currencies acquired are exchanged for domestic currency at these rates, there is no net change in money holdings. What are the values of the other currencies in terms of B's currency? in terms of C's? of D's?

Suppose that the values of the currencies were fixed by an international

monetary authority at 1 of A = ⅓ of B = ¼ of C = ⅙ of D. What will happen to the holdings of the various currencies by the various countries, assuming that all foreign currencies acquired are exchanged for the domestic currency at these rates? How might the various countries react to these changes? Construct new payments tables showing some possible reactions.

2. The following equations give the home absorption and imports functions for countries I and II, with notation as in appendix 2.

$$H_1 = 13 + \frac{4}{15}P_1 \qquad H_2 = 10 + \frac{1}{3}P_2$$

$$M_1 = 2 + \frac{2}{15}P_2 \qquad M_2 = 4 + \frac{1}{10}P_2$$

a. Calculate the equilibrium values for P_1 and P_2.
b. Calculate the equation of the economic warfare curve, and show that it is satisfied by the equilibrium values of P_1 and P_2.
c. Calculate the home absorption multipliers for the two countries (i) under conditions of foreign trade, and (ii) without foreign trade. Check the results by reworking the equilibrium values of P_1 and P_2 on the assumption (i) that H_{10} moves from 13 to 14, and (ii) that H_{20} moves from 10 to 11.
d. Develop formulas for the imports multipliers and check them with the numerical example.
e. Give a graphical solution of the above system, and show that it conforms with the algebraic solution.

CHAPTER 8

DYNAMIC MODELS

THE NATURE OF DYNAMIC SYSTEMS

Up to this point we have dealt mainly with static, or equilibrium, models, with only incidental reference to the dynamic properties of these systems. Every equilibrium model, however, is a special case of a dynamic system, and furthermore there are many dynamic systems which do not lead to any equilibrium position. All systems of the real world are dynamic systems, and very rarely in social or economy systems do we find an actual case of equilibrium in which the variables of the system remain at the same level for successive periods of time. Equilibrium models are valuable mainly because if the equilibrium position is stable, we know that within limits the system will move towards it. Hence, if we are given the equilibrium state of a system and its present state, we can at least predict the direction of change in that the "near" future states will be somewhere between the present state and the equilibrium state. We cannot, however, know for certain whether an equilibrium state is stable or not until we have specified the dynamic properties of the system.

A "state" of a system consists of its description at a moment of time, and a dynamic process consists of a succession of states, like the successive frames on a movie reel. If there are stable relations between states that are a constant "distance apart" in the succession, we have a dynamic system. Thus if "today" bears a constant relation to "yesterday," tomorrow will bear the same relation to today that today does to yesterday. Only if we know the dynamic system underlying a process are we able to predict. If we know a stable relation between yesterday and today, and we know the state of the system today, we can predict the state tomorrow; knowing the state tomorrow we can predict the state the day after tomorrow, and so we can go on indefinitely into the future. Equally, we can

follow the system back indefinitely into the past. Thus, suppose S_t represents the state of the system at any date, t. If a stable relation exists,

$$S_t = F(S_{t-1}) \tag{1}$$

we have a dynamic system of the first degree. Then, if we are given S_o, $S_1 = F(S_0)$, $S_2 = F(S_1)$, and so on. The whole sequence, S_0, S_1,—etc. can be generated from any one state S_t. If a stable relation exists

$$S_t = F(S_{t-1}, S_{t-2}) \tag{2}$$

we have a dynamic system of the second degree. Here today, (S_t), depends not only on yesterday, (S_{t-1}) but also on the day before yesterday, (S_{t-2}). If we are given *two* successive states we can predict all the others of the sequence. Similarly, in a dynamic system of the nth degree, S_t is a stable function of the previous n states, and we must know n successive states before the system can be predicted. An equilibrium system is one in which $S_t = S_{t+1} = S_{t+2} \ldots = S_e$. S_e is the equilibrium value of the system. Thus a dynamic system is in equilibrium if today is such that tomorrow is the same as today.

Difference Equations

Where the state of the system can be described by the value of a single variable, the dynamic system can be described by a simple *difference equation.* A difference equation expresses a stable relationship between two or more successive positions of the same variable. Suppose, for instance, that the value of a loan at interest doubles in a given period, say ten years. Then if x_t is the value of the loan at any date, and x_{t+1} is the value ten years after that date, we can write $x_{t+1} = 2x_t$. This is a very simple difference equation. If we are given the value of x at any date whatever, we can deduce the value at any other date.[1] Thus if $x_0 = 100$, we know from the equation that $x_1 = 200$. But if $x_1 = 200$, then x_2 must be 400. Similarly x_3 is 800, x_4 is 1600, x_5 is 3200, and so on—each value being double the value before. If the value of the variable is initially such that the same value is repeated period after period, the initial value is said to be the *equilibrium* value of the system. That is, if $x_0 = x_1 = x_2, \ldots = x_n$, and so on, $x_0 = x_e$, the equilibrium value. The equilibrium value is found by putting $x_{t+1} = x_t = x_e$ in the difference equation. In the above case the only equilibrium value is $x_e = 0$. This equilibrium is not, however, *stable* in the above case. A slight variation above (or below) zero will set the system moving away from equilibrium. Consider by contrast the system $x_{t+1} = 0.5(x_t)$. If $x_0 = 1$, $x_1 = 0.5$,

[1] Any other date, to be exact, which is an exact multiple of the unit time interval of the equation.

$x_2 = 0.25$, $x_3 = 0.125$, and so on. A disturbance clearly results in a movement back to the equilibrium position at zero.

Application to the Basic National Income Model

These concepts can now be applied immediately to a simple national income system. Suppose we assume a linear consumption function, $C_{t+1} = C_0 + cY_t$. That is, we assume that consumption in any one year depends on the income of the year before. The linear form of the equation means that the graph of this consumption function is a straight line. The propensity to consume, c, is a constant, and it is assumed that there will be some consumption, C_0, even at zero income. Then we suppose investment (willing accumulation) to be constant at A. We know from the savings-investment identity that $Y_{t+1} = C_{t+1} + A$. Putting together the consumption function and the savings-investment identity, we get a difference equation in national income,

$$Y_{t+1} = C_0 + A + cY_t \tag{3}$$

The equilibrium value is found by putting $Y_{t+1} = Y_t = Y_e$, in this equation, giving

$$Y_e = C_0 + A + cY_e, \text{ or } Y_e = \frac{C_0 + A}{1 - c} \tag{4}$$

An arithmetical example will clarify the dynamics of this system. Suppose $C_0 = 40$, $A = 20$, $c = 0.8$. The equilibrium value is then $Y_e = \dfrac{60}{(1 - 0.8)} = 300$. Table 14 then shows first the course of the national income if we start from a point below the equilibrium, say 200, and

TABLE 14. DYNAMICS OF A NATIONAL INCOME MODEL
($Y_{t+1} = 60 + 0.8(Y_t)$)

National Income	Y_0	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9
$C_0 + A$		60	60	60	60	60	60	60	60	60
$0.8(Y_t)$		160	176	189	199	207	214	219	223	226
Y_{t+1}	200	220	236	249	259	267	274	279	283	286
$C_0 + A$		60	60	60	60	60	60	60	60	60
$0.8(Y_t)$		320	304	291	281	273	266	261	257	254
Y_{t+1}	400	380	364	351	341	333	326	321	317	314

then the course of the national income if we start from a point above the equilibrium value, say 400. It will be observed that the equilibrium position is stable. A divergence from equilibrium either above or below

the equilibrium value sets in motion forces to restore the equilibrium position. It will be observed also that there is no cyclical movement. The national income constantly approaches its equilibrium value at a diminishing rate. Fig. 16 shows the path toward an equilibrium of 300 with

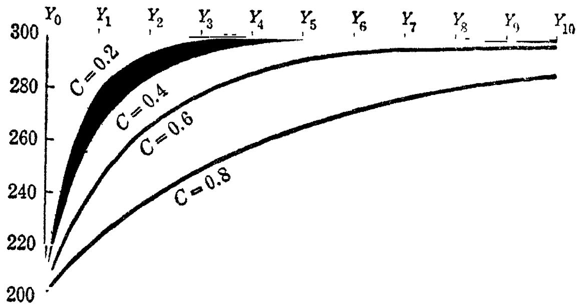


Fig. 16. Paths to Equilibrium

propensities to consume of 0.8, 0.6, 0.4, and 0.2. It is evident that the lower the propensity to consume, the more rapid the approach to equilibrium.

Graphic Solution of Dynamic Models

Systems in which the difference equations relate only two successive years can be "solved" by a simple graphic technique. Suppose we measure x_t on the horizontal and x_{t+1} on the vertical axis in Fig. 17. The difference equation relating these two quantities can be represented as a line on this graph. Suppose in Fig. 17 LK is such a line. Now draw

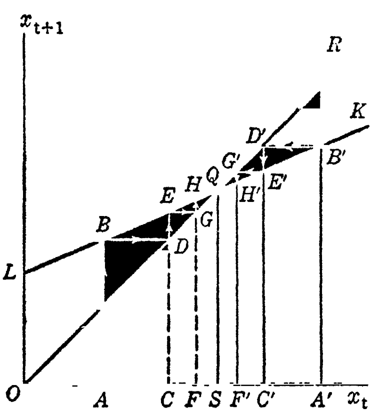


Fig. 17. Solution of a Dynamic Model

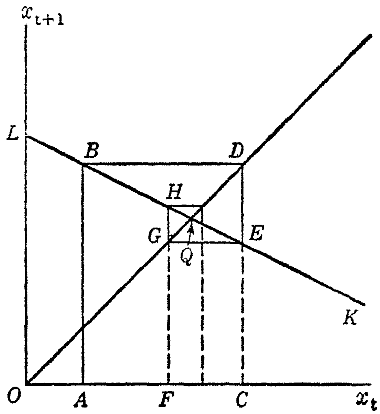


Fig. 18. A Cobweb Model

the 45 degree line from the origin, *OR*, intersecting *HK* in *Q*. *Q* is clearly the equilibrium value, for by construction $OS = SQ$—that is, $x_t = x_{t+1}$. Suppose now that *OA* is the initial value x_0. If *AB* is drawn vertically to meet *HK* in *B*, $AB = x_1$. If *BD* is drawn horizontally to meet *OA* in *D*, we have by construction $OC = CD = AB = x_1$. If *CD* is projected upward to meet *HK* in *E*, $CE = x_2$. Similarly, $FH = x_3$, and so on toward the equilibrium at *Q*. Similarly, starting from a position above the equilibrium at *A′* the successive positions x_0, x_1, x_2, x_3 are *OA′*, *A′B′*, *C′E′*, *F′H′*, etc. The equilibrium in Fig. 17 is clearly stable. If the curve *LK* is steeper than 45 degrees where it cuts the line *OR*, the equilibrium is unstable.

A curve which expresses a difference equation may be called a "difference curve." If the slope of the difference curve is negative, as in Fig. 18, the movement toward or away from equilibrium follows a cyclical path. Starting from an initial position $x_0 = OA$, we have $x_1 = AB(= OC)$, $x_2 = CE(= OF)$, $x_3 = FH$, and so on in a "cobweb" around the equilibrium point, *Q*. The celebrated "cobweb theorem" in price theory is merely a special case of this type of dynamic system. If the inclination of the line *LK* is -45 degrees, *G* will fall on the line *AB*, and the cycle will be repeated indefinitely. If the inclination is numerically greater than 45 degrees the cycle will be an explosive one, and the equilibrium will be unstable.

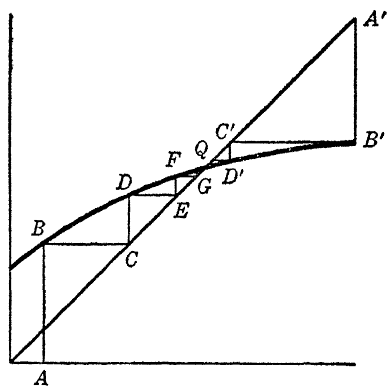


Fig. 19. A Nonlinear Dynamic Model

This graphic technique has the advantage that it can be applied just as easily to nonlinear as to linear difference equations. Thus in Fig. 19 we suppose that the difference curve has a diminishing slope. In terms of our economic model this would mean that the marginal propensity to con-

sume diminished with increasing income—a not implausible hypothesis. Under these circumstances the approach to equilibrium from below, following the course *ABCDEF,* etc., is *slower* than the approach from above, *A'B'C'D',* etc. This would suggest that downswings of the system would be dramatic and rapid, whereas upswings would be slow and prolonged. There is some evidence that business cycles tend to follow this pattern. It is clear also from the analysis that a simple "multiplier" system cannot by itself yield a truly cyclical fluctuation about an equilibrium unless the propensity to absorb is negative—an excessively unlikely hypothesis.

THE ACCELERATOR

If, therefore, there are to be true cycles in a national income model, additional relationships must be introduced. The simplest of these is the "accelerator." This assumes a relationship between the level of investment and the *change* in national income. Let us suppose, for instance, that investment is a constant, K_i plus a constant proportion, α, of the change in national income between the two previous years. That is

$$I_{t+2} = I_0 + \alpha(Y_{t+1} - Y_t) \tag{5}$$

The constant, α, is called the *accelerator,* largely because it reflects the assumption that a rise in income accelerates itself through the stimulation which is given to investment. This stimulation is generally supposed to result from the fact that a rise in output requires at least a proportionate rise in the total stock of equipment necessary to produce this output. If the output of this equipment has been just sufficient in the past to replace a given stock as it wore out, the increase in the stock necessitates a temporary increase in the output of equipment proportionally greater than the increase in final product. Thus, suppose that 1 machine produces 10 widgets, and lasts for 5 years. If the annual output of widgets is 100, there will be 10 machines, 2 of which will be replaced each year. Now suppose that the output of widgets increases to 120. In the year of the increase 12 machines will be needed, so that the output of machines must jump from 2 to 4; 2 for replacement, and 2 more to bring the total number up to 12. A 20 percent increase in the output of widgets, therefore, necessitates a 100 percent increase in the output of machines. This topic will be examined in more detail on pages 197–199.

The Accelerator Model

To return now to the dynamic model, we suppose as before that there is a linear consumption function, so that consumption is a con-

stant, C_0, plus a fixed proportion, β, of the national income of the year before. The constant, β, is the propensity to consume. We have then

$$C_{t+2} = C_0 + \beta Y_{t+1} \tag{6}$$

From the savings-investment identity we derive immediately a difference equation of the second order in national income

$$Y_{t+2} = C_{t+2} + I_{t+2} = C_0 + \beta Y_{t+1} + I_0 + \alpha(Y_{t+1} - Y_t) \tag{7}$$

1. No Cycle

The properties of this system can be illustrated conveniently by arithmetical examples. In all these examples, suppose $C_0 = 800$, $I_0 = 200$, and $\beta = 0.8$, so that the equilibrium value is the same in each, equal to $\frac{1000}{(1-0.8)} = 5000$. Tables 15A to 15E show the course of national income with accelerators, α, equal to 0.1, 0.5, 1.0, 1.1, and 2.0. It will be observed that the influence of a small accelerator is simply to speed up the movement toward equilibrium. Thus the accelerator comes into play in year 2, when it produces a component of investment equal to the rise in income between years 0 and 1 (200, from 4000 to 4200) multiplied by the accelerator, 0.1: $200 \times 0.1 = 20$. Had it not been for the acceleration, national income would have been 4360 in year 2; because of the investment produced by the rise in income, therefore, the rise in income in the next period is greater. That is to say, when income is rising, the accelerator gives it an additional "push." Similarly, when income is falling, the accelerator also gives it an additional push downward and makes it fall faster. In Table 15A this additional push is small enough so that it merely accelerates the movement toward equilibrium.

2. The Damped Cycle

If the accelerator itself is a little larger, however, the push will be great enough so that income will shoot up beyond the equilibrium value of 5000. Once this happens, the consumption function part of the equation (βY_{t+1}) operates to pull the income down again, and a cycle is set up. If the accelerator is less than 1, this cycle will be "damped"—that is to say, the successive fluctuations will make smaller and smaller swings until the swings finally become imperceptible. This is shown in Table 15B, where the accelerator is 0.5. It will be seen that by year 6 the push of the accelerator is sufficient to carry the national income beyond the equilibrium value of 5000, up to 5070 in year 7; in year 8, however, the force of the consumption function is powerful enough to overcome the

TABLE 15A. THE ACCELERATOR MODEL ($\alpha = 0.1$, $\beta = 0.8$)

Year	1	2	3	4	5	6	7	8	9	10	11
C C_o		800	800	800	800	800	800	800	800	800	800
C βY_{t+1}		3200	3360	3504	3618	3706	3774	3826	3866	3897	3921
I I_o		200	200	200	200	200	200	200	200	200	200
I $\alpha(Y_{t+1} - Y_t)$			20	18	14	11	8	6	5	4	3
Y_t	4000	4200	4380	4522	4632	4717	4782	4832	4871	4901	4924
$Y_{t-1} - Y_{t-2}$		200	180	142	110	85	65	50	39	30	23

TABLE 15B. THE ACCELERATOR MODEL ($\alpha = 0.5$, $\beta = 0.8$)

Year	0	1	2	3	4	5	6	7	8	9	10	11	12
C C_o		800	800	800	800	800	800	800	800	800	800	800	800
C βY_{t+1}		3200	3360	3568	3758	3902	3993	4040	4056	4053	4041	4027	4015
I I_o		200	200	200	200	200	200	200	200	200	200	200	200
I $\alpha(Y_{t+1} - Y_t)$			100	130	119	89	57	30	10	−2	−7	−8	−7
Y_t	4000	4200	4460	4698	4877	4991	5050	5070	5066	5051	5034	5019	5008
$Y_{t-1} - Y_{t-2}$		200	260	238	179	114	59	20	−4	−15	−17	−15	−11

TABLE 15C. THE ACCELERATOR MODEL ($\alpha = 1.0$, $\beta = 0.8$)

Year	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C C_o		800	800	800	800	800	800	800	800	800	800	800	800	800	800	800
C βY_{t-1}		3200	3360	3648	4006	4363	4647	4802	4796	4631	4340	3981	3626	3346	3197	3209
I I_o		200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
I $\alpha(Y_{t-1} - Y_{t-2})$			200	360	448	446	355	193	—7	—206	—364	—449	—444	—350	—186	+15
Y_t	4000	4200	4560	5008	5454	5809	6002	5995	5789	5425	4976	4532	4182	3996	4011	4224
$Y_{t-1} - Y_{t-2}$		+200	+360	+448	+446	+355	+193	−7	. −206	−364	−449	−444	−350	−186	+15	+213

TABLE 15D. THE ACCELERATOR MODEL ($\alpha = 1.1$, $\beta = 0.8$)

Year	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C C_o		800	800	800	800	800	800	800	800	800	800	800	800	800	800	800
C βY_{t-1}		3200	3360	3664	4066	4494	4867	5104	5144	4959	4582	4050	3456	2911	2530	2404
I I_o		200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
I $\alpha(Y_{t-1} - Y_{t-2})$			220	418	552	590	513	326	55	−232	−519	−730	−817	−749	−525	−173
Y_t	4000	4200	4580	5082	5618	6084	6380	6430	6199	5727	5063	4320	3639	3162	3005	3231
$Y_{t-1} - Y_{t-2}$		200	380	502	536	466	296	50	−211	−472	−664	−743	−681	−477	−157	+226

weak accelerator, and income starts to fall. By year 15, however, equilibrium is practically reached, and the fluctuations become very small.

3. The Constant Cycle

We contrast this with the case of Table 15C, where the accelerator is equal to 1.0. Here the "push" of the accelerator up when income is rising and the "pull" down when income is falling are so great that a perpetual cycle is set up. The accelerator pushes income up so rapidly that the equilibrium point is overshot by year 3; the consumption function brings the rise to a stop at about 6000 in year 6 and income starts to fall again; the accelerator pushes it past the equilibrium again in year 10, but it reaches a minimum of about 4000 in year 13. Arithmetical roundings cause the figures of the table to be slightly inaccurate; it is clear, however, that income will continue to fluctuate between 4000 and 6000 indefinitely.

4. The Explosive Cycle

If now the accelerator is greater than 1.0, but not too great, explosive cycles will be set up, each fluctuation being larger than the next. This is shown in Table 15D, where the accelerator is 1.1. The accelerator carries income well beyond equilibrium up to a miximum at about 6430, then plunges it down to a minimum at about 3005, far below the original value of 4000. The reader may calculate for himself that from this point income will rise to a maximum of about 7808 before it turns once again for a still more violent fall.

5. The Explosion

Finally, if the accelerator is still larger, the accelerator effect outweighs the consumption function effect altogether, and income moves always in the direction of the accelerator, though perhaps eventually at a decreasing pace; the consumption function is never powerful enough to reverse the direction into a cycle, not even an explosive cycle. This is illustrated in Table 15E, where the accelerator is 2.0. It will be seen that income rises

TABLE 15E. The Accelerator Model ($\alpha = 2.0$, $\beta = 0.8$)

Year		0	1	2	3	4	5	6	7
C	C_0		800	800	800	800	800	800	800
	βY_{t-1}		3200	3360	3808	4742	6462	9410	14222
I	I_0		200	200	200	200	200	200	200
	$\alpha(Y_{t-1} - Y_{t-2})$			400	1120	2336	4300	7368	12032
Y_t		4000	4200	4760	5928	8078	11762	17778	27254
$Y_{t-1} - Y_{t-2}$		200	560	1168	2150	3684	6016	9476	

indefinitely, with the accelerator also rising. These five cases are portrayed graphically in Fig. 20.

Significance of the Models

The economic significance of these dynamic systems is unfortunately much less clear than the mathematical analysis. It must be recalled first

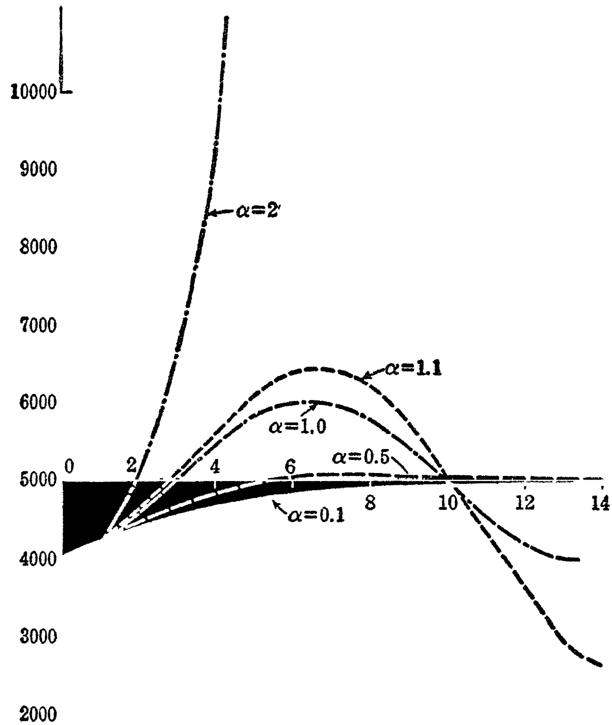


Fig. 20. Movements of National Income with Various Values of the Accelerator (Propensity to Consume = 0.8)

that the development of these models in time assumes constancy in the basic coefficients, the propensity to consume and the accelerator. This assumption is dubious even in the case of the propensity to consume. Forecasts of the postwar national income of the United States made in the later years of World War II generally turned out to be very inaccurate, mainly because they were based on values of the consumption function derived from the 1930s. In the 1940s the consumption function apparently took a marked upward shift, for reasons which are easier to perceive after the event than before. Quite a slight shift in the level of the

consumption function, however, can create large shifts in the position of equilibrium, especially if the propensity to consume is close to 1.0. If the stability of the consumption function is in some doubt, there is much more doubt about the stability of the accelerator. In point of fact there is not much empirical evidence even for the existence of an accelerator, much less for its stability. Hence models which assume constancy in the accelerator must be used with the greatest caution, and they are not likely to be descriptive of actual economic cycles. The best that models of this kind can do is to show how cycles *might* be generated. Whether cycles are in fact generated in the way indicated cannot, of course, be deduced from the theoretical models themselves!

Expectational Price Cycles

Accelerator models showing the course of market prices over time can be constructed along lines similar to those outlined above. Suppose, for instance, that we modify the market identity[2] $p = \frac{Mr_a}{Ar_m}$ into the form $p_t = Kr_t$, where $K \left(= \frac{M}{A} \right)$ is assumed constant and r_t $(= r_a/r_m)$ is a parameter expressive of the intensity of net money demand for the commodity. Let us further suppose that this demand coefficient, r_t, is equal to some normal level, r_0, plus a factor which is larger if the price is rising, smaller if the price is falling. That is to say, suppose that people project the trend of prices, so that rising prices lead to the expectation of further rise, and thus an increase in demand, while falling prices lead to an expectation of further fall, and thus a decrease in demand. If these functions are linear, we can write

$$r_t = r_0 + \alpha(p_t - p_{t-1}) \tag{8}$$

We have, therefore,

$$p_t = Kr_t = Kr_0 + K\alpha(p_t - p_{t-1}),$$

or

$$p_t(1 - K\alpha) = Kr_0 - K\alpha p_{t-1} \tag{9}$$

This is a simple linear difference equation of the first degree; the course of the price through time can therefore be analyzed by a diagram similar to Fig. 18. Thus in Figs. 21A and B the line DD' is the difference curve corresponding to equation (7). If $K\alpha > 1$, the slope of the difference curve DD', $\frac{-K\alpha}{1 - K\alpha}$, will be positive, but less than 1, as in Fig. 21A. The

[2] See Volume I, Chapter 7.

equilibrium at P_e will therefore be unstable. If we start from a price above the equilibrium price, OR_e, say, OP_0, the subsequent prices will be OP_1, OP_2, etc.—the price rises at an accelerating rate. Similarly, if we start at a price below the equilibrium the price will fall at an accelerating rate. As $K\alpha$ gets closer to 1, the line DD' becomes steeper, and the movement

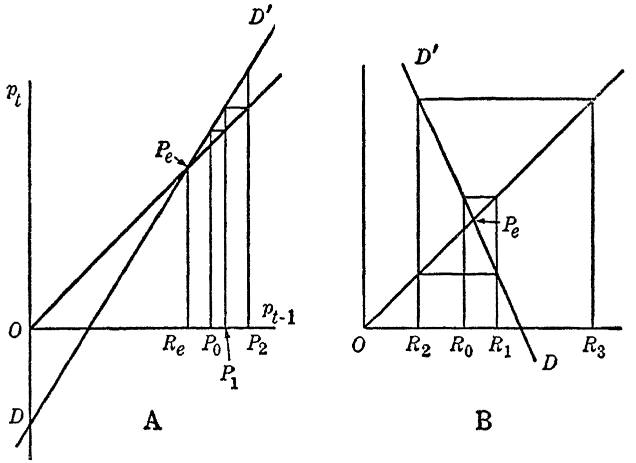


Fig. 21. Price Cycles

away from equilibrium gets faster and faster, until when $K\alpha = 1$, DD' is vertical, and the equilibrium is perfectly unstable; the price shoots immediately to plus or minus infinity if disturbed from the equilibrium level. With $K\alpha < 1$ a cyclical pattern will be set up, as in Fig. 21B. If $K\alpha$ lies between 0.5 and 1, the cycle will be explosive, as in the figure; the course from OR_0 is up to OR_1, down beyond OR_0 to OR_2, up beyond OR_1 to OR_3, and so on. If $K\alpha = 0.5$, the cycle will be perpetual. If $K\alpha$ lies between 0 and 0.5, the cycle will be damped, as in Fig. 18. The smaller $K\alpha$, the more rapidly is the cycle damped. If $K\alpha$ is zero, the price moves immediately to its equilibrium position, $p_e = Kr_0$.

If α is negative, so that rising prices lead to the expectation of a fall, DD' will have a positive slope less than 1, that is, less than 45 degrees, and the price will move steadily toward the equilibrium, as in Fig. 17.

The abrupt transition from high to low to high prices in Fig. 21B is not particularly realistic. However, this is merely the result of expressing the model in the form of a difference equation of the first degree. If we suppose that r_t is a function of $(p_{t-1} - p_{t-2})$, say,

$$r_t = r_0 + \alpha(p_{t-1} - p_{t-2}) \tag{10}$$

we get a difference equation of the second degree,

$$p_t = Kr_0 + K\alpha(p_{t-1} - p_{t-2}) \tag{11}$$

The reader can test for himself that this relationship also yields a perpetual cycle if $K\alpha = 1$, a damped cycle if $K\alpha < 1$, and an explosive cycle if $K\alpha > 1$.

POPULATION ANALYSIS

A very important field of dynamic analysis covers the theory of the movements of a population. A population may be defined as any aggregate which can be measured in some common unit, and in which the *age* of each unit can be identified, or at least postulated. The age of a unit is the length of time which has elapsed between some "present" date and the date of "birth," or the entry of the unit into the population. In all ordinary populations it is also generally postulated that each unit has a finite age of "death," or departure from the population. Birth and death in this sense do not necessarily mean the creation or the destruction of the unit in question, but refer to the dates at which the unit begins to conform to, or ceases to conform to, the definition of the aggregate which constitutes the population. Thus, if our aggregate is the human population of a finite area, such as the United States, the entrance of an immigrant is a "birth" and the exit of an emigrant a "death" into that population, just as much as physiological births and deaths, with one important difference: natural births result in individuals of zero age, whereas immigrants may add members of any human age to the population.

The Survival Function

The dynamic analysis of population movements rests on the postulate of certain stable difference equations. We may, for instance, postulate a stable "survival function," showing how many units survive to each age from a given number of births. Thus, of a 100 births, 90 may survive to be one year old, 85 to be two years old, 81 to be three years old, and so on down to the last tottering survivors. The survival function is thus built up from a series of difference equations,

$$a_1 = k_1 a_0,\ a_2 = k_2 a_0,\ \ldots\ a_n = k_n a_0 \tag{12}$$

where a_n is the number that survive to year n out of a total number of births, B_0, in year 0. In the simplest form of the function the survival coefficients, $k_1, k_2 \ldots k_n$, are constants. The number dying, or leaving the population, out of the group that is r years old in the year t, between the year t and the year $t + 1$ is $a_r - a_{r+1}$, or $k_r a_0 - k_{r+1} a_0$, or $\left(\dfrac{k_r - k_{r-1}}{k_r}\right)\alpha_r$. The survival function can therefore also be expressed as "death function."

Dynamics of a Stable Population

If now we are also given a "birth function" relating the number of births in any given year to the numbers or composition of the population, the whole course of the population can be charted. The process can conveniently be illustrated by arithmetical examples. Consider, for example, the process shown in Table 16. Each row in this table represents the numbers in a population divided into five age groups. Thus suppose that in year 1 there are 100 units between the ages of 0 and 1 year, 96 between 1 and 2, 66 between 2 and 3, 56 between 3 and 4, 30 between 4 and 5, and none older than 5 years. The diagonals marked by the arrows show the survival distributions of each "cohort," or group in the population having the same birth year. Thus we suppose that of 100 people between the ages of 0 and 1 at the end of year 1, 80 survive to the end of year 2, 60 to the end of year 3, 40 to the end of year 4, 20 to the end of year 5, and none to the end of year 6. If the "net births," B, in each year are defined as the number between 0 and 1 year old at the end of the year, then we suppose that the survival function in the above table is given by $k_1 = 1$, $k_2 = 0.8$, $k_3 = 0.6$, $k_4 = 0.4$, $k_5 = 0.2$, $k_6 = 0$. The figures in parentheses at the top of the table show the net births from which the age groups of year 1 are descended, if the survival function is assumed to be constant. Thus, if there are 96 units between 1 and 2 years old in year 1, these must be the survivors of a cohort of 120 births in the previous year: 96 = 120(0.8). In year 2, then, the numbers between 2 and 3 years old must be 72 [= 120(0.6)] and in year 3 the numbers between 3 and 4 must be 48 [= 120(0.4)] and so on.

TABLE 16. POPULATION DYNAMICS: STABLE POPULATION

Year	Age (Years) 0–1	1–2 (120)	2–3 (110)	3–4 (140)	4–5 (150)	5+	Total
1	100	96	66	56	30	0	348
2	100	80	72	44	28	0	324
3	100	80	60	48	22	0	310
4	100	80	60	40	24	0	304
5	100	80	60	40	20	0	300
6	100	80	60	40	20	0	300

The Death Function

The same process can also be described by using a "death function." Thus, if a number $(0.6)B$ aged 2–3 becomes $(0.4)B$ by the end of the following year, $(0.2)B$ have died during that year—that is, one-third of the number aged 2–3. Thus, of the 66 aged 2–3 in year 1, 22 will die in the following year, leaving 44 survivors aged 3–4 in year 2.

In Table 16 we have supposed that the number of net births after year 1 is constant at 100 per year. It is evident on this assumption that the population soon settles down to an equilibrium value of 300, in which the same age grouping is repeated year after year. An equilibrium population can be defined as one in which the number of births and deaths per annum is equal and constant, and in which the numbers in each age group are also constant.

Population Dynamics: Increasing Population

Suppose now that the number of births each year is not constant, but is a function of the numbers in the different age groups. As long as this function is known and is stable it is still possible to trace out the course of any population through time. Thus suppose, in Table 16, that following the year 6 the number of births in each year was equal to the total number of units in the 2–3 and 3–4 age groups. In year 7 the number of births would still be 100 [$= 60 + 40$] and the population would continue in equilibrium. Suppose, however, that the number of births was equal to the number of units in the 1–2 and 2–3 age groups. The population would follow a course as in Table 17. In this table the first figure in each row (after the first) is equal to the sum of the 1–2 and 2–3 age groups of the preceding year. Thus $140 = 80 + 60$, $172 = 112 + 60$, and so on. It is

TABLE 17. POPULATION DYNAMICS: EXPANDING POPULATION

Year	Age (Years) 0–1	1–2	2–3	3–4	4–5	5+	Total
6	100	80	60	40	20	0	300
7	140	80	60	40	20	0	340
8	140	112	60	40	20	0	372
9	172	112	84	40	20	0	428
10	196	138	84	56	20	0	494
11	222	157	103	56	28	0	566

assumed that all those born in the course of a year are between 0 and 1 in the following year. The survival function is assumed to be the same as in Table 17. It is evident that on these assumptions the population will increase indefinitely. Every increase in births provides a bigger age group as a base for still more births.

Declining Population

Suppose, on the other hand, we assume that the number of births in each year is equal to the numbers in the 3–4 and 4–5 age groups. The population will follow the course of Table 18. In this case each cohort

TABLE 18. POPULATION DYNAMICS: DECLINING POPULATION

Year	Age (Years) 0–1	1–2	2–3	3–4	4–5	5+	Total
6	100	80	60	40	20	0	300
7	60	80	60	40	20	0	260
8	60	48	60	40	20	0	228
9	60	48	36	40	20	0	204
10	60	48	36	24	20	0	188
11	44	48	36	24	12	0	164
12	36	35	36	24	12	0	143

of births produces less than its own number to replace it, and the population will decline indefinitely.

Algebraic Solution

It is not difficult to formulate the conditions under which populations will expand, contract, or be stationary under stable birth or survival functions. Let $a_1, a_2, \ldots a_n$ be the numbers in the n age groups of a population. Then the birth function can be described by a series of constants, $b_1, b_2, \ldots b_n$, so that

$$B = b_1 a_1 + b_2 a_2, \ldots b_n a_n \tag{13}$$

This is the simplest useful form of the function. Some of the b's may of course be zero, and in a biological population they will be largest in the most fertile age groups. The complication introduced by the necessity for the cooperation of two sexes is neglected here. Now we also postulate a survival function, also as a set of constants, $k_1, k_2, \ldots k_n$, as before. We have, therefore, the year t,

$$a_1 = k_1 B_{t-1},\ a_2 = k_2 B_{t-2}, \cdot \cdot \cdot\ a_n = k_n B_{t-n} \tag{14}$$

It follows, combining equations (13) and (14), that

$$B_t = b_1k_1B_{t-1} + b_2k_2B_{t-2}, \cdots + b_nk_nB_{t-n} \tag{15}$$

If now the population is to be stable, $B_t = B_{t-1} = \ldots = B_{t-n}$, and

$$R = b_1k_1 + b_2k_2 + \ldots + b_nk_n = 1 \tag{16}$$

The Net Reproductive Ratio

The quantity, R, is a measure of the net reproductive ratio. In an equilibrium population the survival distribution of each cohort of births is the same as the age distribution in any year. We see this in Table 16, where it is clear that the survival distribution, once the population has settled down to equilibrium (the series on the marked diagonals), is the same as the age distribution of the rows: 100, 80, 60, 40, 20, 0. In an equilibrium population, therefore, R measures the number of net births to which each cohort of the population gives rise. If $R = 1$, therefore, the population just reproduces itself: each generation as it dies out leaves a new generation of equal size. If the net reproductive ratio is more than 1 the population will increase indefinitely, as in Table 17, for each generation as it dies out leaves behind it a cohort of larger size than itself. Similarly, if $R < 1$, the population will eventually decline indefinitely, for each generation leaves behind a cohort smaller than itself.

The Dynamics of Constant Growth Rates

A dynamic equilibrium of population at a constant rate of growth is possible, though only under rather restricted conditions. Consider, for instance, the population of three age groups given in Table 19.

The basic difference equations behind this table are derived from the coefficients of the survival distribution, 1, 0.8, and 0.4, and the coefficients of the birth function, $B_t = 3a_2 + 8a_3$. Thus in year 1 the number of

TABLE 19. POPULATION IN DYNAMIC EQUILIBRIUM WITH A CONSTANT RATE OF GROWTH

Year	Age (Years) 0–1	1–2	2–3	Total
1	100	40	10	150
2	200	80	20	300
3	400	160	40	600
4	800	320	80	1200

births is $40 \times 3 + 10 \times 8$, or 200; of these 200, 200 survive to be the first age group of year 2, 160 survive to be the second age group of year 3, and 80 survive to be the third age group of year 4. It is clear from the table that the population grows at a constant rate of 2 per annum, and that the number of births and the numbers in each age group grow at the same rate.[3]

CAPITAL GOODS AS A POPULATION

Population analysis is important not only for the light which it sheds on the development of human populations, but because it can be applied to populations of all kinds, animate and inanimate. In economics, the analysis is of particular importance in the theory of capital because capital goods of all kinds can be regarded as populations, provided that a survival function can be postulated. This means that it must be reasonable to suppose stable, or moderately stable, relationships between the *age* of goods and their rates of consumption or survival. In most cases this assumption is by no means unreasonable. Thus, automobiles form a population, in which the age of each automobile can be determined with some accuracy and in which "life tables" or "survival tables" can be constructed showing what proportion of the automobiles "born" or produced in any one year will survive into other years, or will "die" (i.e., be scrapped) in other years.

Birth Functions of Capital

Populations of goods differ from biological populations mainly in the totally different nature of their birth functions. In biological populations, as we have seen, it is reasonable to suppose that the number of births in any year is a function of the numbers in the various age groups of the population. Automobiles do not beget automobiles, however, in the way that horses beget horses; the genetic processes of physical capital are complicated, and involve practically the whole economic system. Nevertheless,

[3] Suppose the equilibrium population grows at the rate, g. Then we must have $B_{t+1} = gB_t$, whence $B_{t+r} = g^r B_t$. Inserting the appropriate values in equation (15) we have (canceling B_t)

$$\frac{b_1 k_1}{g} + \frac{b_2 k_2}{g^2} + \ldots + \frac{b_n k_n}{g^n} = 1 \tag{17}$$

Thus, given all the b's and k's, the corresponding equilibrium rate of growth (or decline) can be derived from equation (17). It is clear from equation (17) that as $g \gtreqless 1$, $R \gtreqless 1$. It should be observed that the existence of an equilibrium rate of growth does not necessarily mean that this equilibrium situation is stable; the system, starting from any arbitrary age distribution, may move toward equilibrium or it may not. The conditions of equilibrium, however, are too complex for the present work.

simple birth functions can be postulated in the case of physical capital, yielding models of population movement which throw a great deal of light on the processes of the real world. Suppose, for instance, that "births" (production) of any good fulfill two functions: one to replace those items which have "died" (been consumed) during the period, and the other to add to the total stock in response to an increased demand. The first may be called "replacement production," the second "expansion production." Total production in any year will be the sum of these two quantities, the second of which, of course, will be negative in the case of diminished demand. If the demand for the good is assumed constant, so that the total stock is maintained at a constant level, the number of births in any year will be equal to the number of deaths in that year; that is, production will equal consumption.

The "One-Hoss-Shay" Cycle

Table 20 shows an interesting model of a population of "one-hoss shays"[4]—that is, goods all of which survive to a given length of life (five years in the table) and then disintegrate. In terms of the preceding analy-

TABLE 20. POPULATION ANALYSIS: ONE-HOSS-SHAY CYCLES

	Age (Years)							
Year	0–1	1–2	2–3	3–4	4–5	5–6	Total	Production
0								100
1	100						100	100
2	100	100					200	100
3	100	100	100				300	100
4	100	100	100	100			400	100
5	100	100	100	100	100		500	100
6	100	100	100	100	100	0	500	200
7	200	100	100	100	100	0	600	100
8	100	200	100	100	100	0	600	100
9	100	100	200	100	100	0	600	100
10	100	100	100	200	100	0	600	100
11	100	100	100	100	200	0	600	200
12	200	100	100	100	100	0	600	100
13	100	200	100	100	100	0	600	100

[4] The reference is to a poem by Oliver Wendell Holmes.

sis, the survival function is given by $k_1 = k_2 = k_3 = k_4 = k_5 = 1$, $k_6 = 0$. Suppose that production is 100 units a year for the first six years. It is evident that by the fifth year an equilibrium population has been reached, and that if the total population is maintained at 500 units, this equilibrium will continue indefinitely, with 100 units dying and 100 being born every year. Now, however, we suppose an increase in demand for the item in the year 6, leading to an increase in the total population from 500 to 600. This can be achieved only by increasing production in that year from 100 to 200. This cohort of 200 gradually passes through the age distribution, being 0–1 year old in year 7, 1–2 years old in year 8, and so on, until it finally dies in year 11. In each year from 7 to 10, 100 units die every year and, as the population is just maintanied in numbers, 100 must be born. In year 11, however, 200 die, so 200 must be produced. This clearly starts the process all over again, and the cycle will be repeated indefinitely, 200 being produced every fifth year and 100 in the intervening years. Similar cycles will be set up wherever there is a distortion in the age distribution, whatever the cause.

Distortions Follow Too Rapid Increase

It should be observed that the distortion follows from the attempt to increase the total stock too quickly. If instead of producing 200 in year 6, 120 were produced in each year thenceforward, a new position of equilibrium would be reached in year 11 which would be permanent, and there would be no cycle established. This is illustrated in Table 21. The number in each age group is increased by an equal amount, and the relative age

TABLE 21. POPULATION ANALYSIS: ONE-HOSS SHAYS WITH STEADY GROWTH

	Age (Years)							
Year	0–1	1–2	2–3	3–4	4–5	5–6	Total	Production
6	100	100	100	100	100	0	500	120
7	120	100	100	100	100	0	520	120
8	120	120	100	100	100	0	540	120
9	120	120	120	100	100	0	560	120
10	120	120	120	120	100	0	580	120
11	120	120	120	120	120	0	600	120
12	120	120	120	120	120	0	600	120

distribution (the proportion of the total stock in each age group) is the same as before. Hence no cycle arises.

Damped Oscillations

These conclusions are substantially modified if we relax the assumption that the commodity is a "one-hoss shay," lasting for a certain number of years and then immediately disappearing from the scene. Just as in

human life not everybody lives to the allotted span and some exceed it, so in the case of a commodity some pass out of existence after a short time while some last longer. Automobiles, like humans, have a life table; some crack up in the first year of their existence, some are still on the road after 20 years. It is not difficult to show that the result of this modification of the original assumption is to give us not a perpetual cycle but a *damped* oscillation—i.e., a cycle in which the amplitude continually diminishes as time goes on. If there were no further initial distortions, then the age distribution would gradually work itself into an equilibrium again. Actually, of course, there are likely to be continual new distortions, giving rise to new cycles.

The Acceleration Principle

The effect of distortions in the age distribution of goods is accentuated by another principle which forms the basis of the "acceleration principle" noted earler. If one commodity, *B*, is necessary for the production of another commodity, *A*, then the fluctuations in the output of *A* will be reflected by intensified fluctuations in the production of *B*. Let us refer again to the previous example. Let us suppose that to produce 10 units of this commodity we must have a machine, *B*. Then, from Table 20,

TABLE 22. POPULATION ANALYSIS: THE ACCELERATOR

Year	Age (Years) 0–1	1–2	2–3	3–4	4–5	5–6	Total	(Idle)	Production
4	2	2	2	2	2	0	10	(0)	2
5	2	2	2	2	2	0	10	(0)	12
6	12	2	2	2	2	0	20	(0)	0
7	0	12	2	2	2	0	18	(8)	0
8	0	0	12	2	2	0	16	(6)	0
9	0	0	0	12	2	0	14	(4)	0
10	0	0	0	0	12	0	12	(2)	20
11	20	0	0	0	0	0	20	(0)	0
12	0	20	0	0	0	0	20	(10)	0

when 100 units of *A* are being produced, a stock of 10 machines is necessary. When 200 are produced in year 6, however, 20 machines are necessary. If we suppose that the machine likewise has a life of 5 years and that it had an equilibrium age distribution, it is easy to see in Table 22 that there will be a very large distortion in its age distribution and a consequent recurrent cycle of very large amplitude. In year 4 everything is in equilibrium. Then in year 5 the number of machines must be increased from 10 to 20, in anticipation of the increase in production of commodity *A* from 100 to 200 units in year 6. Production in that year must, therefore, be 12 machines—2 for replacements and 10 to expand the number from

10 to 20. Then in year 6 the number of machines is 20, but there is no need for production at all, as in year 7 only 10 machines will be required to produce the 100 units of commodity *A*. Hence, there will be no production of machines in years 7 to 10. In year 7 there will be 18 machines, with 8 idle. In year 8 there will be 16 machines, 2 having disintegrated, with 6 idle, and so on to year 10, when 20 machines must be produced, as the 12 that were produced in year 6 disappear, leaving none. In year 11, 20 machines are required to produce the 200 units of commodity *A*. Here again a perpetual cycle is set up with 20 machines produced every five years, none at all produced in the intervening years, and an average of 4 machines idle. This is as extreme as a cycle can be. If now we assume that the production of commodity *A* is spaced out as in Table 21, so that no distortion in its age distribution results, there will still be a distortion in the age distribution of the machine, *B*, as shown in Table 23. Now a

TABLE 23. POPULATION ANALYSIS: THE SECOND ACCELERATOR

Year	Age (Years) 0–1	1–2	2–3	3–4	4–5	Total Stock	Production
4	2	2	2	2	2	10	2
5	2	2	2	2	2	10	4
6	4	2	2	2	2	12	2
7	2	4	2	2	2	12	2
8	2	2	4	2	2	12	2
9	2	2	2	4	2	12	2
10	2	2	2	2	4	12	4
11	4	2	2	2	2	12	2
12	2	4	2	2	2	12	2

stock of 12 machines is required in each year to produce the 120 units of commodity *A*. In year 5, therefore, the production of machines will increase from 2 to 4 units, and this will be repeated every five years. In this case the fluctuation is not so intense as in Table 22. Nevertheless, if we now suppose that a machine tool, *C*, is necessary to produce machine *B*, the fluctuations in the output of the machine tool will be much greater than the fluctuations in output of the machine. Again, it must be remembered that the severity of the fluctuations in the above examples are due to taking the very extreme case of a "one-hoss shay" life table. In fact the smoother life tables of reality operate to dampen and smooth out these oscillations. Nevertheless, it remains true even in the general case that fluctuations in the output of machine tools are greater than fluctuations in the output of machines, and fluctuations in the output of machines are greater than the fluctuations in the output of the commodities which they make. This principle is called the acceleration principle because the demand for machines depends not so much on the demand

for their product as upon the *rate of change* in the demand for the product. It is clear that the problem of adjusting the rates of growth of the stock of various kinds of goods so that distortions in the age distributions are avoided is practically insoluble. In this sense, therefore, some fluctuation in economic activity is a necessary cost of economic progress, a cost which will be greater the more rapidly we attempt to advance.

Distortions in Age Distribution

Both in the case of human populations and populations of goods, distortions in age distribution from any cause may take a long time to work out. The sudden fall in infant mortality in many tropical countries about 1950 is creating a youth problem in the 1960s; even in the United States the rise in the birth rate in the 1940s is creating a problem of unemployment among youths in the 1960s, as unusual numbers of young people come on to a labor market that is adjusted to smaller numbers of entrants. The fact that virtually no automobiles were produced in the years 1942-45 in the United States because of the war created some difficulties in Detroit in 1958-60, when there were no automobiles in the age groups of heaviest scrapping rate.[5] Age distribution distortions sometimes produce what is called a "shadow effect" in the second generation; thus, the high birth rates of the 1940s in the United States produce high marriage rates and high birth rates in the 1960s and 1970s, when the children of the 1940s are of childbearing age.

The "Austrian" Theory of Capital

The analysis of capital as a population of goods is of importance in interpreting the "Austrian" theory of capital, associated mainly with the name of Böhm-Bawerk.[6] According to this view, capital is created when "original factors of production" (labor and land) are embodied in goods, and is destroyed when the services of these embodied factors are finally realized. Capital, therefore, is regarded as a "population of value," into which values are "born" when factors of production are employed and "die" when services are yielded up. The "average period of production" then is the average "age at death" of values—i.e., the average interval of time which elapses between the employment of a factor of production and its final consummation as utility. Thus the bread that I eat today represents services of delivery men this morning, bakers yesterday, millers last month, farmers last year, and so on. The average age of the services embodied in the bread is its average period of production.

[5] See K. E. Boulding, "An Application of Population Analysis to the Automobile Population of the United States," *Kyklos*, Vol. 2 (1955), pp. 109–124.

[6] E. von Böhm-Bawerk, *The Positive Theory of Capital*, trans. by William Smart, London, 1891; Stechert, 1923.

In a population in stationary equilibrium the average age at death is equal to the ratio of the total population to the annual number of births or deaths. Thus, if 100 units are born into and die out of a population each year, and the average age at death is 30 years, the total population will be 3000.[7] In the stationary state, therefore, the ratio of total capital (the total population of value) to the annual income (income of factors [births] or consumption of product [deaths], these being equal) is equal to the average period of production (the average length of life of embodied values). If the average period of production lengthens, this must mean either an increase in the total capital, if income is constant, or a decrease in income, if capital is constant. Attempts have been made (e.g., by Hayek)[8] to interpret depressions in terms of a decline in income necessitated by an overexpansion of the period of production with an inadequate capital stock. The possibility of difficulties of this kind cannot be altogether ruled out. However, ordinary business-cycle depressions seem to be of a totally different nature. The basic difficulty with period of production analysis is that it is strictly valid only in a stationary state, and is too crude an analytical device to deal with the complex motions of dynamic systems.

DYNAMICS AS ADJUSTMENT TO STRESS

The difference equation models of this chapter are useful in calling attention to certain possible patterns in the dynamic systems of the economy. We should be careful, however, not to put too much weight

[7] Suppose an equilibrium population with age groups $a_1, a_2, \ldots a_n$. Suppose that B is the annual number of births or deaths, and all deaths occur at the end of the year. Then $a_1 - a_2$ dies at age 1, $a_2 - a_3$ at age 2, $a_{n-1} - a_n$ at age n. The average age at death then is

$$T = \frac{(a_1 - a_2) + 2(a_2 - a_3) + 3\,(a_3 \; - a_4) \ldots + n(a_n)}{B}$$
$$= \frac{a_1 + a_2 \ldots + a_n}{B} = \frac{P}{B}$$

In dynamic equilibrium with a constant rate of growth, g, if $a_1, a_2, \ldots a_n$ are initial age groups, the survival distribution for the group a_1 is $a_1, ga_2, g^2a_3, \ldots g^{n-1}a_n$. The average age at death then is

$$T = \frac{(a_1 - ga_2) + 2(ga_2 - g^2a_3) + 3(g^2a_3 - g^4a_4) \ldots + ng^{n-1}a_n}{B}$$
$$= \frac{a_1 + ga_2 + g^2a_3 \ldots + g^{n-1}a_n}{B}$$

In this case $T \gtreqless \frac{P}{B}$ according as $g \gtreqless 1$, though both T and $\frac{P}{B}$ are constants.

[8] F. A. Hayek, *Prices and Production,* London, Routledge, 1931.

on them as predictors, for the parameters of social systems can change in quite unpredictable ways. Thus, in the 1940s many predictions of future population were made that turned out to be highly inaccurate, mainly because of a quite unexpected rise in fertility. Furthermore, purely mechanical dynamic systems, which are very successful in describing and predicting the solar system, where difference (or differential) equations of the second or third degree are all that are necessary, are quite inadequate to describe the complex processes of social systems, where current behavior may depend on memories and records going back many years into the past. Human history, indeed, is a dynamic system of almost infinite degree and, because of this, has a large element of unpredictability. It is not, however, wholly unpredictable, and analysis can, at least, give us some conditional predictions of the form "if we do A, B is likely to follow."

Furthermore, the general dynamic process of society can usually be interpreted in terms of adjustments to stress. Here we see that mechanical dynamics does not displace equilibrium analysis in so far as "stress" can be interpreted as a divergence between the current state of a system and its equilibrium state. Up to a point stress may produce no change; the resistance to a change is too great for the stress to overcome. At some point, however, the stress overcomes the resistance and change occurs. By and large, the greater the resistances to change, the larger the changes when they finally come. Such systems are sometimes called "threshhold" systems, as the stress must reach a certain threshhold before any change takes place. Once the threshhold has been reached, some adjustment of the system must be made, and the principle here is that what adjusts is the adjustable; that is, those parts of the system which have the lowest threshholds will adjust first.

We have seen many examples of this principle. Thus, in a competitive market, prices are highly adjustable, and any strain in the form of shortages or surpluses will shift prices. In monopolistic or organized markets, prices are not easily adjustable, and strains may produce adjustments in preferences or technology. An income deflation will produce a fall in prices and wages where these are adjustable and a fall in output and employment where prices and wages are not adjustable. Pressure to change sometimes produces increased resistance to change, and sometimes a decreased resistance. In the light of these complexities and uncertainties it is clear that mechanical dynamic models should be treated as illustrations rather than as predictors.

QUESTIONS AND EXERCISES

1. "The ability to predict future events in any science depends on its ability to discover stable difference equations relating the variables of its universe."

Why, exactly? Does this proposition throw any light on the relative success of astronomy, meteorology, and the social sciences in their attempts at prediction?

2. It is argued that without an analysis of the dynamic properties of an equilibrium system it is impossible to judge the *stability* of the equilibrium. Why?
3. It has been suggested that economics should confine its analysis of change to "comparative statics"—that is, the comparison of two equilibrium positions in a model, on the grounds that the actual dynamic processes of the economy involve variables which belong more to the sociologist or the psychologist, such as the speed of rumor, or the velocity of responses to stimuli. Is this a possible explanation of the alleged unsatisfactory nature of dynamic economic models? Is the general suggestion valid?
4. In some older analyses of business cycles much attention is paid to various "leads" and "lags" in the time series of the economy—e.g., a lag of wages behind prices. How would such leads or lags be expressed in terms of difference equations? Experiment with the construction of systems of such difference equations and explore their properties.
5. Expectations play an important role in some types of dynamic analysis. Why? What sort of links might be postulated between present expectations and past experience? Could these links be described in terms of difference equations? If a link can be established between present expectations and present behavior, can we then link present behavior directly with past experience without any specific postulation of expectations? Does this mean that dynamic systems are possible without introducing expectations?
6. Consider the following dynamic model:

$$C_t = 600 + 0.7\,Y_{t-1}$$
$$I_t = 100 + 0.1\,Y_{t-1}$$
$$C_t + I_t = Y_t$$

C_t = consumption, I_t = investment, Y_t = income in year t.

What is the equilibrium value of *C, I, and Y?*

Starting with $Y_0 = 3000$, trace the course of the various variables for a period of ten years. Repeat, starting with $Y_0 = 4000$. Use both the arithmetic and the graphic method.

7. Consider the following dynamic model:

$$C_t = 600 + 0.7Y_{t-1} + 0.1K_{t-1}$$
$$I_t = 100 + 0.1Y_{t-1} - 0.1K_{t-1}$$
$$C_t + I_t = Y_t$$
$$K_t = K_{t-1} + I_{t-1}$$

K_t is the total capital stock of the society in year t. The other symbols have the same meaning as in Exercise 6.

Comment on the economic significance of these equations.

What is the equilibrium value of the variables of this system?

Starting with $K_0 = 4000$ and $Y_0 = 3000$, trace out the course of the various variables for a few years.

Can you devise a graphic solution to this system?

Repeat the exercise, substituting

$$C_t = 600 + 0.7Y_{t-1} + 0.2K_{t-1}$$

for the first equation.

What light does this model throw on the "stagnation thesis"?, that is, the view that a mature capitalism cannot achieve full employment because of lack of investment opportunity.

8. Consider the explosive model of Table 15E, page 185, $Y_t = 1000 + 0.8Y_{t-1} + 2(Y_{t-1} - Y_{t-2})$. Suppose now that the economy reaches capacity output at a level of 7000 so that it is impossible to raise output above that level. Suppose also that there is a floor of -1000 below which investment cannot fall, for physical reasons. Trace the course of national income under these restrictions for 20 years. Note that the "booms" are short and the "depressions" long. What feature of the model produces this result? How realistic is such a model?
9. Repeat Exercise 8, with the *additional* assumption that consumption also has a floor of 3000 below which it cannot fall.
10. A population of four age groups has a survival function $k_1 = 1$, $k_2 = 0.9$, $k_3 = 0.5$, $k_4 = 0.1$, and a birth function $b_1 = 0$, $b_2 = 2$, $b_3 = b_4 = 0$.
    a. Start with an age grouping $a_1 = 180$, $a_2 = 121$, $a_3 = 50$, $a_4 = 7$ in year 0, and trace the course of the population for ten years. At what rate is the population growing in each year?
    b. Repeat, with an initial age grouping of $a_1 = 100$, $a_2 = 10$, $a_3 = 10$, $a_4 = 10$. Note that the population exhibits a two-year cycle. Why? Is this a realistic model for human or animal populations?
11. A population of four age groups has a survival function $k_1 = 1$, $k_2 = 0.9$, $k_3 = 0.8$, $k_4 = 0.7$. The birth function is $B_t = D_t$, where D_t is the number of deaths in year t.
    a. Start with an age grouping $a_1 = 100$, $a_2 = 90$, $a_3 = 80$, $a_4 = 70$, the total population being 340. Trace the course of the population for a few years and show that the population is in equilibrium. (Note: You will find it useful to insert a line for "deaths" between the age grouping of each year.)
    b. Suppose now that the total population is suddenly increased to 440 by an additional 100 births. Trace the course of the population for about 20 years, and show that it exhibits a damped oscillation in the age groups. Why?
    c. Repeat the above exercise with a survival function of $k_1 = 1$, $k_2 = 0.5$, $k_3 = 0.2$, $k_4 = 0.1$, and an initial age grouping $a_1 = 100$, $a_2 = 50$, $a_3 = 20$, $a_4 = 10$. Note that the damping of the oscillation is much more pronounced. What feature of the model leads to this effect?

CHAPTER 9

ECONOMIC FLUCTUATIONS AND GROWTH

THE NATURE OF CYCLES

One of the most striking phenomena of the Western world is the alternation of periods of prosperity and depression which goes by the name of the business cycle. This is perhaps the most complex problem of modern economic society, and we cannot hope to explore all its ramifications in a single chapter. With the analytical tools now at our disposal, however, it is possible to outline the principal forces at work and to suggest certain remedies.

Is There a Cycle?

It must be emphasized at the outset that the business cycle is an extremely complex phenomenon. Indeed; it has been questioned whether there is any such thing as "a" cycle—whether the so-called business cycle is not merely the result of a combination of accidental irregularities with certain special cycles in particular industries. It is an instructive exercise to construct a completely random time series—e.g., by throwing dice or pennies, picking numbers out of a hat, or counting the frequency of occurrence of a letter in a book. The eye that is accustomed to tracing cycles in the time series of economic life will almost immediately begin to detect evidence of cyclical movements in any purely random series. If we are to prove that there is a business cycle, therefore, we must do something more than show that there are fluctuations in economic quantities, such as prices or employment. These fluctuations must be shown also to have some degree of regularity and a fairly regular period of recurrence. Perhaps the best evidence of the existence of nonrandom fluctuations in economic life, however, is to be found in the connections that are to be

observed between the movements of different economic series. Thus, it is observable in the record that low prices generally go hand in hand with depression and unemployment; that a fall in prices is usually followed by a fall in wages after a certain interval, and so on. These indications support the belief that economic fluctuations are not random, but follow a more or less regular pattern.

True and False Cycles

It is still somewhat open to question, however, whether this regular pattern can be strictly be defined as a *cycle.* A true cyclical movement is a swing about a position of equilibrium, in which there are forces operating to bring the fluctuating quantity back to the equilibrium once it diverges from it, and in which the momentum of a movement towards the equilibrium point carries the quantity beyond it. The pendulum is the simplest example of such a cyclical movement: the further the bob swings from the equilibrium position, the stronger is the force operating to bring it back, but the very swing that brings it back gives it momentum that sends it swinging to the other side. In a frictionless world this could go on forever. There are undoubtedly some economic cycles (e.g., the hog cycle) of this nature. The business cycle cannot simply be interpreted, however, as the swing of a pendulum. There are some movements in economic life which tend to perpetuate themselves indefinitely unless checked by some outside force. The process of monetary inflation and deflation are somewhat of this nature. Hence, there may be cyclical movements which are more like the oscillation of a billiard ball between two cushions than the swing of a pendulum—in which a certain movement might go on indefinitely until reversed by some outside influence.

Endogenous and Exogenous Changes

The question must be raised also to what extent the business cycle is the result of accidental historical disturbances, such as inventions, wars, explorations, and discoveries. These external disturbances produce great fluctuations in the economic system, and yet there may be nothing in their essential nature to make them regular. That is to say, the business cycle may be an illusion created by the fact that in the past two hundred years there have been great external disturbances which happen to have come along at moderately regular intervals. This view regards the business cycle largely as an accident of economic development, or rather as a result of the inevitable irregularity of economic development. There is much to be said for this view, especially as it does not necessarily exclude the possibility of genuinely cyclical changes superimposed on the unfolding process of economic history. Indeed, the problem of the business

cycle cannot be separated from the theory of economic development—i.e., the study of the time relationships of economic quantities. Economic history, like all history, may be regarded as an amalgam of freedom and necessity. Certain causal factors can be traced through the whole stream of history—A causing B; B, C; C, D; and so on indefinitely. Into this causal nexus there continually strike new factors—original events which arise out of freedom, chance, call it what you will. So in economic development we distinguish between "endogenous" changes—changes which are generated within the economic system itself—and "exogenous" changes, which intrude into the economic system from outside, such as political or technological events that do not seem to be closely related to economic processes. The question whether in the whole course of historical events there are any exogenous changes is a difficult philosophical question which fortunately we do not have to answer. The fact that we can divide events, however roughly, into "economic" and "noneconomic" is sufficient justification for postulating the existence of exogenous changes as far as economic life is concerned.

Economic and Noneconomic Events

The distinction between what is and what is not economic is by no means easy to draw. In so far as all decisions are matters of choice they may be said to be governed by economic principles. Nevertheless, there is an important, if unclear, distinction between those choices which are made consciously and voluntarily, with weighing up of advantages and disadvantages, and those which are made involuntarily, without conscious weighing of alternatives, or which are motivated by the unregulated passions. Thus, the decision of an American middle-class family to have a child is economic in the sense that the decision of an Indian peasant is not. Similarly, many political decisions, wars, etc., are made without any clear weighing of alternatives, with the passions of love and fear, envy and greed, as the main motivating factors, and without any adequate process of testing the reality of our images of the alternatives among which we might choose.

A "Standard" Cycle

Whatever doubt there may be about the regularity or causation of the cycle, it is possible to describe five states or conditions of the economic system which, if put together in order, would constitute a "standard" cycle.

1. Depression. There is first a state of depression, characterized by low output and underemployment of both men and equipment. Prices and

wages are likely to be low relative to a condition of boom; but perhaps even more important, the relative structure of prices is badly distorted. The price of finished products is likely to be low relative to the price of labor; the real wages of those employed may be high, for though money wages are low, the price of things which workers buy—wage goods—are lower still. The distribution of the national income is sharply distorted. Profits are low or even negative. Interest, because it tends to be fairly constant in dollar terms in short periods, becomes a much larger proportion of national income. The proportion of national income going to wages and salaries is likewise high, in spite of the widespread unemployment. Because of the fall in the absolute real total of the national income, however, the total of real wages and salaries is likely to be less than in the previous period of prosperity. The high real wages of those who are employed are more than counterbalanced by the poverty of the mass of unemployed, and even in the case of the employed the fear of unemployment takes the edge off their greater purchasing power.

The price of raw materials and agricultural commodities is also likely to be low relative to the price of industrial goods. The "terms of trade" between manufacturing and agriculture have moved in favor of manufacturing, but the advantage which this would otherwise give to the manufacturing population is eaten away by the fact that manufacturing output and employment are so low. Agricultural output, however, will be maintained at a fairly high level. Money incomes will be low generally, but the reason for these low money incomes is different in different sectors of the economy; in agriculture, employment and output will not have fallen, and may even have risen above the boom level. The low incomes of the agricultural population are due to the low prices of agricultural commodities. In industry, prices do not fall so much, and the low incomes are due to low output and unemployment. The industries particularly affected by unemployment are the construction and durable goods industries—building, steel, automobile, and so on. Those least affected are the nondurable consumption goods industries—food, clothing, and the like. There is very little construction, either of building or machines.

Net investment is low, or even may be negative in extreme cases. Velocity of circulation is low, bank loans and deposits are low, stock prices are low, but earnings are also low because of low profits and low dividends. Bank reserve ratios may be high, for banks cannot find enough borrowers.

2. *Recovery.* A depression is generally followed by a period of recovery. Employment and output rise, unemployment falls. With the rise

in output there is, of course, a rise in real income. In the early stages of recovery, prices are usually fairly stable, so that there is a rise in money income proportionate to the rise in real income. As a proportion of national income, profits rise, interest and wages fall, though the total wage bill in real terms may well increase because the rise in total income outweighs the fall in labor's proportionate share. Investment rises, perhaps in part spontaneously, in part under the influence of rising profits. Bank loans and deposits rise, though rates of interest may remain low or even continue to fall. The *proportion* of national income going to consumption falls because of the rise in investment, but the absolute amount of real consumption probably rises, again due to the rise in real national income. The velocity of circulation of money probably increases, leading to a larger volume of payments to finance the increased money incomes. The durable-goods industries revive—construction, machinery, machine tools, automobile, durable household goods. The acceleration principle, noted earlier (page 181), may come into play, leading to still greater growth in investment and in income. Stocks of farm products, raw materials, and inventories of goods of all kinds may diminish; there may be a certain flow of money from household into business balances. Farm prices, and therefore farm incomes, rise with the recovery of industrial employment and output. The whole process may be cumulative. On the monetary side increased expenditures lead to increased receipts, which in turn may lead to increased expenditures, and so on. On the "real" side increased employment leads to increased incomes and increased consumption, which induces further investment and so leads to still larger income. This process may, however, stop short of full employment, or it may go on until full or optimum levels of employment and output are reached.

Recovery periods differ considerably in the forces which initiate or sustain them. The impetus may come from innovations, from the exploitation of new fields for investment in new lands, new products, or new methods. The impetus may come merely from the need to replace capital which has been consumed or depreciated during the depression. Or the impetus may come from government expenditure, for instance, on armaments or war. Whatever the impetus, however, the general dynamics of a recovery is likely to follow the lines indicated above.

3. Full Employment. The state of full employment, or prosperity, is presumably the most desirable state and an important goal of economic policy. There are no resources that are *involuntarily* unemployed, the owners of which would like to employ them at prevailing prices but can find no employers. In the field of labor this means that at prevailing wages anyone who wants to work could find a job after a reasonable interval of time. Even in a state of full employment there will be, of course,

a certain amount of unemployment caused by labor turnover—i.e., by the fact that it takes a certain amount of time to pass from one job to another; but this is voluntary in the sense that people are willing to take this risk if they quit the job they are now performing. Output will be at the highest level that could be permanently maintained. A higher level of output than this is possible temporarily, but only at the cost of hidden capital consumption or an overworked labor force. Money income will be approximately stable, or may be rising slowly in proportion to the rise in real income brought about by economic progress. Prices will also be approximately stable. Costs and receipts, and therefore profits, are at the level that will just call forth an amount of enterprise sufficient to employ the voluntarily employable population. The structure of relative prices will be approximately "normal"—e.g., the relation of agricultural to industrial prices will be such that resources are being driven from agriculture at a rate commensurate with the requirements imposed by the existing rate of technical progress.

4. Boom. This halcyon state is not necessarily reached in all its phases during a recovery, and even if reached, is not likely to be stable. Recovery frequently leads not to a stable state of full employment, but to a boom. The rise in expenditures leads to a rise in prices, once full employment approaches in various industries and supply curves become inelastic and begin to shift. This rise in prices, if it leads to expectations of a further rise, may lead, as we have seen, to a flight from money into any and every kind of real capital. Buildings go up everywhere; new factories, new offices, appear; many new businesses are started. Full employment may pass over into "hyperemployment"—i.e., a shortage of labor, where instead of more workers than there are jobs, there are more jobs than there are workers to fill them. Money wages rise, but prices rise faster, and real wages fall. Profits are abnormally high, and hence seem to justify the mass of new investment. There is a boom on the stock market; everybody is anxious to get rid of money and hold stocks, so the price of stocks rises rapidly. Banks make loans to almost anyone who wants them, and people spend the money so obtained in buying commodities, securities, or labor.

5. Recession. The boom, however, is unstable, and tends to pass into a recession. It becomes apparent after a while that some of the enterprises which have been started in the rosy anticipations of the boom are not going to be profitable. Or it becomes evident that the opportunities for investment are dwindling. When three skyscrapers have been built in the old home town it becomes evident that a fourth would not only be highly unprofitable but would ruin the other three. Indeed, it may be even that the third, which is almost complete, is already seen to be one too many.

Consequently, building slows down, unemployment appears in the construction industries, expenditure declines, and incomes decline with it. This means that other enterprises cease to be profitable. Prices begin to fall, and consumers' purchases—and therefore the revenues of businesses—decline. Real wages frequently stay up at first, as profits decline sharply. Many businesses shut down or contract their operations, and unemployment becomes general, leading to further decline in incomes, further retrenchment of expenditures, and more unemployment. Meanwhile, in the stock market the boom will have to come to an end; the price of stocks will begin to fall, creating a flight *into* money, a strong desire to sell stocks, and a catastrophic fall in stock prices. Banks get scared, and try to improve their liquidity position by calling in loans and selling investments. This, however, reduces the volume of deposits, and leads to a further fall in prices; the fall in the prices of assets may soon wipe out the whole of a bank's net worth and cause a wave of bank failures. This leads to further restriction of expenditure, more unemployment, an almost complete cessation of investment, and we are back again in a depression, whence, like a continuous performance at the movies, the show goes on again; only, unlike the movies, we have to sit through it.

Irregularity of Phases

While there is an observed tendency for economic fluctuations to pass successively through the phases outlined above, the movement may be highly irregular and one phase does not necessarily lead immediately into the next, nor is the length of each phase at all definite. Thus a state of fairly stable depression may lead into recovery, or it may decline into a further recession, as in Britain in 1929. Similarly, a recovery may level off, or even turn into a recession, far short of full employment, and even further short of a boom, as in the United States in 1937. Depression may be fairly stable, and recovery slow, as in the 1930s, or depression may be unstable and recovery rapid, as in 1921. Prosperity may be fairly stable, as from 1924 to 1929; or it may be highly unstable, as in 1919. We may have a rather stable mild depression, as in the late 1950s and early 1960s. Nevertheless, even though the periods and the order of the various elements of fluctuation are highly variable, they represent recognizable conditions of the economic system.

THE ANALYSIS OF BUSINESS CYCLES

It would be impossible in a single chapter to give more than a very partial account of the methods of analysis of business cycles, especially as much of the ground has already been covered in previous chapters. What we can do here perhaps is to draw together some of the methods

and the problems which arise out of them in a broad conspectus which will show us the wood rather than the trees.

Dangers in Eliminating the Trend

We should first perhaps warn the reader against too great reliance on any single analytical method, for every method inevitably involves half-hidden assumptions which may distort the picture of reality in the mind of the investigator. Consider first some of the problems which arise in the mere description of business cycles. It is not within the scope of this work to describe the various statistical methods employed. Nevertheless, we should be on our guard to see that statistical methods do not imply, and indeed create, a picture of reality which may be misleading. A frequent prerequisite for the statistical analysis of cycles, for instance, is the elimination of the trend. The method of elimination is not important; what we must realize, however, is that in eliminating the trend we are assuming that there is one—that is, we interpret the cycle itself in terms of *fluctuations about* or deviations from some "norm" which may be a pure figment of the statistical imagination.

The Trend Values May Not Be "Normal"

There are two reasons for supposing that statistical trend lines may be quite misleading. In the first place the significant *deviation* of the observed variable may not be from any median values, but from some extreme values. Suppose, for instance, that in Fig. 22A, we measure time along OT and some economic variable—say employment—along OA. The irregular line shows the course of the variable through time. Suppose now the dotted line represents what would be "full employment" at each time, as measured, say, by the total labor force. Unemployment is then seen to be always a deviation *below* the norm. Suppose, however, that the statistician persuades us to draw a statistical trend line through the data, as in Fig. 22B. The deviations then appear partly above and partly below the statistical "trend." In this case, however, the statistical trend may be economically meaningless, because there is an economically significant deviation which is not described by the statistical deviations. The picture of the cycle which is given in Fig. 22B as a series of fluctuations. alternating between "too much" or "too little" employment may be quite misleading, if in fact all deviations, or what is even worse, *most* deviations are really below the norm. It thus makes a great deal of difference to our picture of the problem itself whether we regard the trend as a clothesline from which all the deviations hang downward, as in Fig. 22A, or whether we regard it as a skewer with the actual values lying about as much above as below the line, as in Fig. 22B.

Trends and Growth Curves

A second objection to the elimination of the trend may be almost more fundamental, and is illustrated in Fig. 23. In Fig. 23A we have drawn a time series which consists of three successive growth curves, P_0P_1, P_1P_2, and P_2P_3. The assumption here is that the variable measured along OA (suppose this is national income) receives some kind of an impetus for growth at P_0—let us say, a new invention. A growth impetus is almost invariably followed by an "ogive" growth curve with the general shape of P_0P_1. Growth is apt to be slow at first, for the impetus takes time to

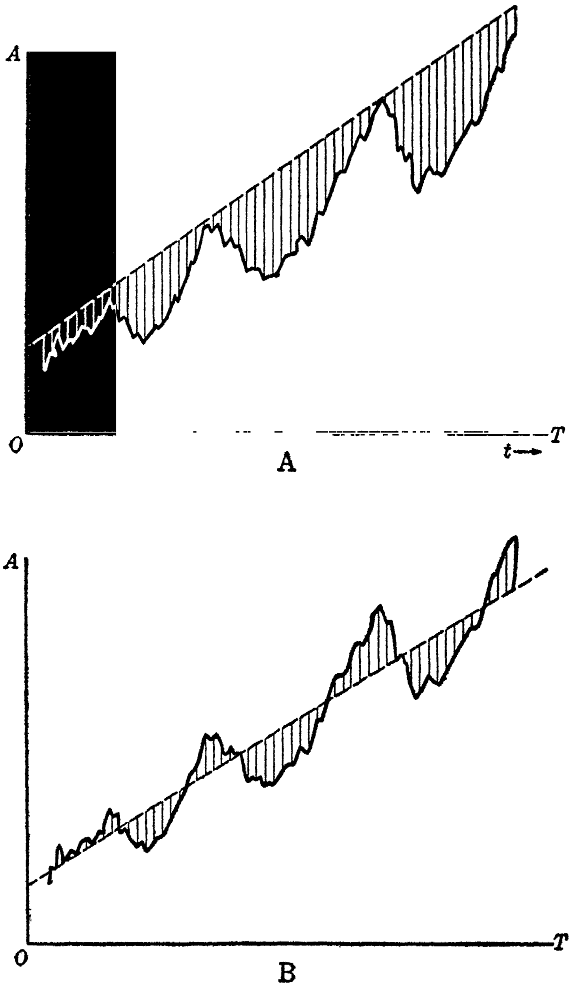


Fig. 22. Trends and Cycles

gather weight. After the period of incubation, however, growth becomes rapid, reaching a maximum rate at about Q_0, after which the rate declines as the original impetus works itself out and becomes exhausted. We then suppose a new impetus at P_1, followed by a new growth curve, $P_1Q_1P_2$, and a third impetus at P_2 followed by a new curve, $P_2Q_2P_3$. Now suppose that the statisticians persuade us to "analyze the cycle" by drawing a

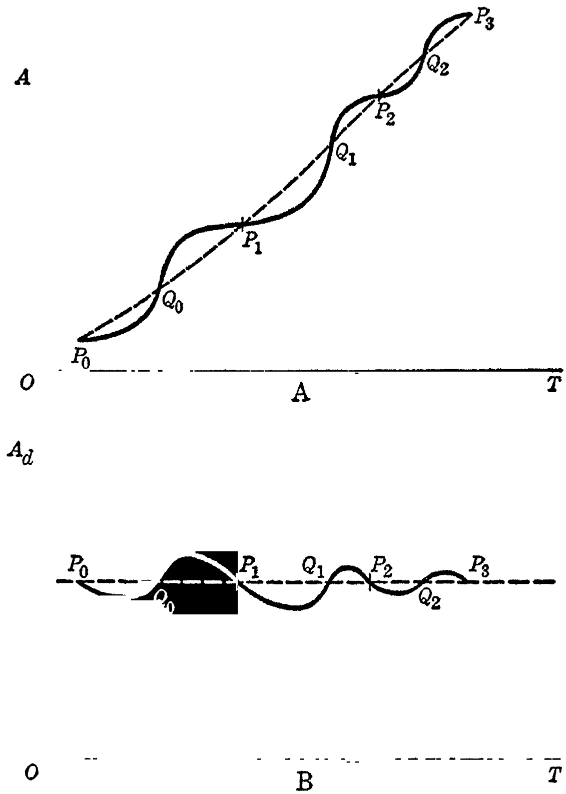


Fig. 23. Trends and Growth Curves

trend line (dotted in the figure) and "eliminating the trend" as in Fig. 23B, where we plot the deviations from the trend line. Then to make matters worse we forget all about the original data and concentrate all our efforts on the analysis of the data as presented in Fig. 23B. It is evident that we have completely changed our picture of the problem—perhaps unknowingly—by our statistical techniques. Growth and the growth curves have disappeared; we are analyzing something that looks like a cycle, and may in fact be nothing of the sort. In this case the growth curve itself is the significant trend; and while it might be important to

discuss deviations from the growth curve, the deviations of the growth curve itself from an arbitrary statistical trend line may be meaningless, or even positively misleading. Thus the point Q_0 in the growth curve actually represents the peak of the growth process, beyond which the forces making for exhaustion of the growth impulse become increasingly dominant. In Fig. 23B, however, the point Q_0 apparently represents the beginning of the boom. Similarly, the point P_1 in the growth curve represents the start of a new impulse—this is the real point of "recovery" from the doldrums of stagnation. In Fig. 23B, however, the same point P_1 indicates the start of a "depression." These examples indicate, perhaps, how important it is to make statistics our servant and not our master!

Dangers in Cyclical Models

Turning now to the dangers which beset the economic analyst as well as the statistical analyst in this field, we may raise the warning here also that it is possible to get absorbed in models of cycles to the exclusion of models of growth. It is easy, as we have seen in the previous chapter, to set up dynamic models of economic systems which contain cycles. We have seen how difference equation systems even of the first order can produce cycles which may be either explosive, stable, or damped. The "cobweb theorem" (pages 179, 188) is an important example.[1] The more complex our models, the more variables we have, and the greater the degree of our difference equations, the more difficult it becomes *not* to get cycles out of the models we construct. We have seen some examples in the previous chapter: the acceleration-multiplier model, for instance (pages 182–185), and the population process models (pages 195–199). The very multiplicity of these models, however, means that they must be used with great care, especially in view of the rather arbitrary assumptions which underly them all. The multiplier-accelerator model for instance assumes stability in both the consumption function and the accelerator coefficient —conditions which are certainly not likely to be true of any actual situation for very long. Population analyses depend on the assumption of stable survival and birth functions, again representing a condition which may only obtain for short periods. The rout of the population forecasters in the late 1940s and 1950s, when the actual human population growth was almost universally far in excess of what was predicted by all the experts, and of the economic forecasters who predicted large-scale depression following the end of World War II, are examples of the dangers of model building based on insecure assumptions in regard to the stability of the various functions and coefficients involved.

[1] See also Volume I, Chapter 12.

The "Sensitivity" (Shiftability) of Models

Nevertheless, some important things can be learned from the economic models. All the macroeconomic models are characterized by fairly high degrees of what may be called "sensitivity." The sensitivity of a model may be defined as the change in its *equilibrium* position which results from a small unit change in one or another of the basic parameters of its behavior equations. Thus, the multiplier is a measure of the sensitivity of the basic model with respect to changes in any parameter which expresses the "height" of the total absorption function, and, as we have seen on page 57, the closer the marginal propensity to absorb draws toward 1, the larger will be the multiplier. In a system with a large multiplier, then, very slight changes in the underlying behavior functions—for example, in the height of the consumption function or the investment function—will produce large changes in the equilibrium output. This is what we mean when we say that the system is sensitive. Generally speaking, the sensitivity of a system depends on the difference between the slopes of the basic equations. This is seen clearly in Fig. 8, page 55, where we see that as the slope of the total absorption function approximates that of the basic identity (unity), the multiplier gets larger and the system grows more sensitive. In general, systems where the equations of equilibrium can be reduced to two equations whose slopes are nearly equal will be sensitive; if the slopes diverge, and still more if one is positive and the other negative, the system will be insensitive.

One of the important discoveries of macroeconomics is that while price equilibria tend to be insensitive, except in the case of industries of decreasing supply price,[2] because the slopes of the demand and supply curves are opposite in sign, income equilibria are highly sensitive for the reasons outlined above. The more sensitive a system is, however, the more likely is it to be subject to large fluctuations in its equilibrium position because of quite small fluctuations in the basic parameters which determine it. Even small and irregular random fluctuations in the basic parameters will lead to an apparently cyclical fluctuation of much greater magnitude in the observed variables because of the time lags involved in moving from one position of equilibrium to another.

A Cycle in the Basic Model

This proposition is illustrated for the basic model in Fig. 24. Let us suppose a stable consumption function, $C'C$, all the fluctuations being the result of shifts in the level of investment between the two extremes, OI_0 and OI_n. The equilibrium position will shift between E_0 and E_n, de-

[2] See Volume I, Chapter 14.

pending on the level of investment. Suppose then that the system is in actual equilibrium at E_0, and there is a sudden shift to the lower investment level. The system will move toward the lower equilibrium, quickly at first and then more slowly as time goes on, following a path such as $E_0E_1E_2$, etc. Before the system actually gets to E_n, suppose the investment level shifts to OI_0 again. The system moves toward the upper equilibrium,

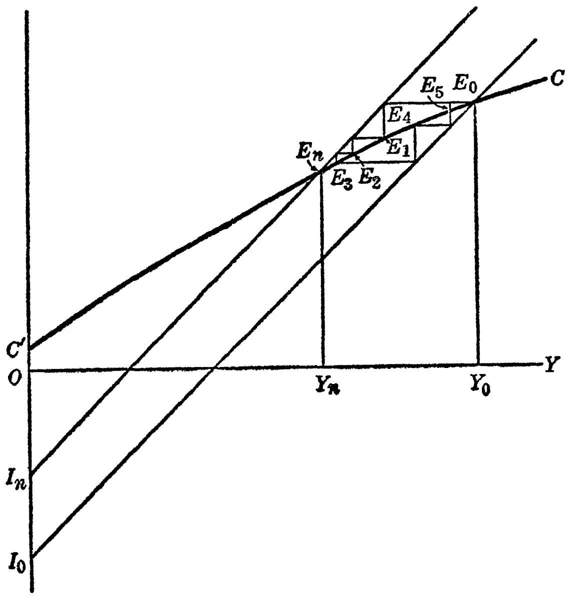


Fig. 24. A Cycle in the Basic Model

again quickly at first, then more slowly, following the course $E_3E_4E_5$, etc. Mere sensitivity will not, of course, lead to regular cycles unless there are regular cycles in the basic parameters. The observed business cycle, however, is not particularly regular, and much in the observed data can be interpreted in terms of random fluctuations in the basic parameters of a sensitive system.

Sources of Sensitivity: Elastic Money Supply

One of the tasks of economic analysis in this regard, then, is to look round the system for possible sources of sensitivity. The most obvious, and probably the most important, we have already discussed in the form of the multiplier. There are, however, many other sources which might be briefly mentioned. If the money supply is elastic with regard to the price level, we may have a monetary equilibrium that is highly

sensitive, as we see in Fig. 10, page 79. If the increase in the money stock which follows a unit change in the price level is approximately equal to the ratio M/P (or T/V), assuming this to be stable, slight changes in the underlying parameters—for instance, in the height of the money supply curve, or in the velocity of circulation—will cause large changes in the equilibrium values of M and P. This is the basic assumption of the Wicksellian system, (pages 141–142) in which slight changes in the parameters due, say, to interest rate changes, produce indefinitely large shifts in the equilibrium position, so that we always have either a perpetual inflation or a perpetual deflation until the parameters change. This model is helpful also in explaining hyperinflation—the condition of runaway price level which occasionally follows war or revolution, and is mainly an outcome of the breakdown of the tax system. We will return to this question in the next chapter.

Profits-Investment Relation

Another possible source of sensitivity in the system is the reciprocal relationship which may obtain between profits and investment If we are going to make a model with an investment function, it would seem reasonable to suppose a behavior relationship between investment and profits. Investment is undertaken mainly in the expectation of profits, and if the expectation of profits is high when profits of the past year have been high, a positive relation may be postulated between profits of one year and investment of the next. There is, however, also another possible relationship between investment and profits: it is not only true that profits create investment; it is also true that, in part at any rate, investment creates profits. The profit-making process is that by which the net worth of a business increases. An increase in the value of the real assets of businesses, however, unless it is offset by a decline in business money holdings or a rise in liabilities, will increase the net worth of businesses and so create profits.[3] If both the investment-profits and the profit-investment relations are stable, we can construct a model for the determination of both profits and investment, illustrated in Fig. 25. In this figure HEH' represents the profits-investment relationship. That is, it shows what amount of investment will be planned for at each level of profit. It is assumed that there must be some profits before any investment is undertaken, and that from the point of zero investment, H, higher profits bring higher investment, but at a diminishing rate. Similarly, the curve KEK' is the investment-profits curve, showing what level of profits is created by each level of investment.

[3] For a further discussion of this problem see K. E. Boulding, *A Reconstruction of Economics,* 2nd ed., New York, Wiley, 1962, chapter 14 and appendix.

It is assumed that profits will be positive (= OK) even if investment is zero, as there are other sources of profit besides investment; then as investment rises, profits also rise, but as there are likely to be increasing "leakages" it is assumed that profits rise at a decreasing rate. It is where investment is self-financed that it is most powerful in creating profits. We have here two opposing forces; the higher the level of investment the

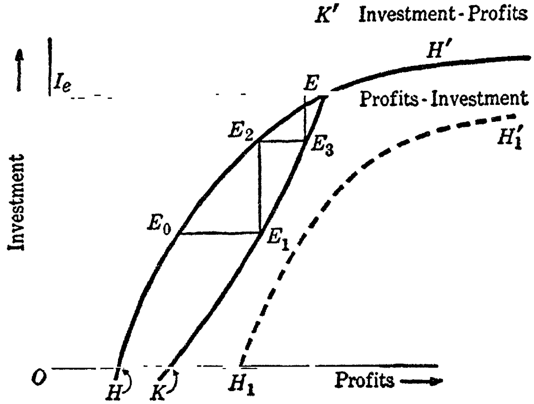


Fig. 25. Dynamics of Profits and Investment

lower the proportion that is likely to be self-financed; on the other hand, the higher the level of profits the more likely is investment to be self-financed. We suppose here that the first of these two forces is more persistent and dominates at higher levels of investment.

The point of equilibrium is where the two curves intersect at E. This equilibrium is stable, as will be seen by the dynamic process; if we start from a position E_0, the amount of investment corresponding to E_0 creates profits corresponding to E_1, which induces investment corresponding to E_2, which creates profits corresponding to E_3, and so we move toward the equilibrium point. If the curves intersected in the reverse direction, the equilibrium would, of course, be unstable. Even if there is a position of stable equilibrium, however, the fact that the curves slope the same way means that the system is highly sensitive to change in one or in both of the basic behavior equations. If the profits-investment curve falls—i.e., if businessmen become more pessimistic, or less willing to invest at each level of profit—the equilibrium will move down the curve EK, both profits and investment being less than before. If the slopes of the curves are nearly equal, a small "worsening" of the basic behavior equations may mean a large decline in both investment and profits. If now we link this figure with Fig. 24, we see that the decline in investment is multi-

plied (by the multiplier!) in its effect on equilibrium output or income, so in combination we have a system quite frightening in its possible sensitivity, where a very slight change in the profits-investment function may make all the difference between full employment and severe depression.

Indeed, a slight shift of the profits-investment curve to the right—say to $H_1H'_1$—would result in a situation with no position of equilibrium, where profits and investment would shrink together in a dizzying spiral of ever-deepening depression. Something like this may have happened between 1930 and 1932.

The Boomerang

Sensitivity, of course, merely explains why little fluctuations in one thing can give rise to big fluctuations in something else. It does not in itself go very far toward explaining the initial fluctuations. We must, therefore, look for another property of economic systems to explain the more or less regular changes in direction of movement which occur in them. This property we might call the "boomerang." This is the element in a system which tends to *reverse* any present direction of movement when that movement has gone far enough. Any system with a boomerang in it will have fluctuations. A movement in any direction will be eventually reversed, and will become a movement in the opposite direction, which in turn will be reversed into a movement in the first direction—this successive reversal of direction being what is meant by a fluctuation.

Continuous Boomerangs

There are at least two different kinds of boomerangs, continuous and discontinuous. The continuous boomerang is the true cycle, such as the pendulum in physics or the accelerator-multiplier system in economics. It is characteristic of the dynamics of an equilibrium system in which the movement toward equilibrium produces acceleration which drives the system beyond equilibrium, but in which the further the system diverges from equilibrium, the stronger is the force operating in the direction of equilibrium. There are many examples of such systems in economics, and the key to their analysis is the search for accelerators. The investment-accelerator has already been discussed. The "price-expectations" cycle is an example of the same phenomenon, the accelerator in that case being the movement of liquidity preference or in velocity of circulation in response to expectations projecting current movements of prices. The banking and financial system likewise provides many examples of accelerators. An inflationary movement of the economy is apt to be reinforced by the expansion of bank loans and of the deposits created thereby, and a deflationary movement is likewise likely to be reinforced by the contraction of bank loans and the resulting shrinkage in deposits. If the

deflationary movement goes far enough to cause a wave of failures among banks, it is accelerated even further. Not the banking system alone but the whole credit and financial system is capable of acceleration effects. The expansion of consumers' credit, for instance, in an inflationary period helps to prolong and accentuate the movement, and its subsequent contraction likewise accelerates the downward movement.

Finance as an Accelerator

The financial system also operates as an accelerator in the investment cycle. When credit is easy, it is easy to "finance"—i.e., spread the ownership—of new accumulations of real capital. Entrepreneurs borrow to build buildings, install machinery, or build up stocks of raw materials or half-finished goods. The "real ownership" of these things lies with the lenders and may be very widely distributed; hence it is relatively easy to absorb large stocks of new goods without deflationary pressure. The existence of easy credit, both from banks and from other sources, enables businessmen to create and to hold the title to goods of long life-period which they do not intend to own permanently. A builder, for instance, will build apartment houses with the aid of credit, in the expectation of selling the buildings to investors when they are completed. During the process of construction, the builder holds the title to the values being created, while the lender of credit may be the owner of the equity which the buildings represent. If the expectation is fulfilled and the builder sells the property, he can pay his debt and go on to build more buildings. But if he finds that there is no demand for his buildings on the part of investors, he will be forced either to own and operate the buildings himself, thus changing the character of his business, or, if he wishes to continue his building business, to obtain more capital. If he cannot do that, and if his creditors, who have usually lent on short terms, insist on being repaid, the title of the building may pass to the creditors. All these events are symptoms of crisis. Indeed, it is not too much to say that a boom is a period when bad investments are made, and a crisis is the period when they are found out. It has been argued with some force that the creation of credit makes it easier to make bad investments, for it prolongs the period during which a bad investment can proceed undiscovered. By enabling businessmen to produce in anticipation of demand, it also enables them to make more mistakes.

The Discontinuous Boomerang

The crisislike character of many of the turning points in economic fluctuations suggests that the model of a discontinuous boomerang may also be relevant. This is the "floor and ceiling" phenomenon, where a

movement is reversed not because it has passed its equilibrium position, but because it hits a boundary of some sort which cannot be exceeded. Thus, we might have a system in which the position of equilibrium lay either at infinity or well outside the limits of the fluctuations, and in which movement proceeded indefinitely until it was stopped by some boundary from which it rebounds—the cessation of movement in one direction producing a reversal. Thus we might have systems of such great sensitivity that the equilibrium position always lies outside the possible boundaries of the variables, and movement, therefore, cannot be reversed by passing beyond the equilibrium position, but can only be reversed by a shock which shifts the equilibrium position of the system from very high to very low values, or the reverse. Thus we may have continuous movements of inflation or deflation which never reach an equilibrium, but which can be reversed by a shock, as in the Wicksellian system (see page 141, and Exercises 8 and 9, page 203).

Floors and Ceilings

An important model involving a system of this kind has been developed in recent years. Suppose, for instance, that, for reasons we shall examine in a moment, full employment can only be maintained if output and investment grow at a certain rate. If this rate is faster than the growth of capacity of the system, it can only be maintained if we start from a period of unemployment. Growth at the full-employment rate, however, cannot then be maintained for very long without the system hitting a ceiling, when full employment is reached, and the rate of growth must then slacken. The slackening of the rate of growth, however, produces a situation in which there is not enough investment to maintain full employment, and a decline follows, which may be accelerated by the various accelerators mentioned above. Decline also cannot go on forever—it will eventually reach a floor. The exact position of the floor is somewhat uncertain, but at least output cannot go below what is necessary to provide subsistence for the system! Once the decline has reached the floor, however, it ceases, the accelerators all cease to function and the possibility of growth at the full-employment rate again appears, and the rise starts over again.

GROWTH, STAGNATION, AND DEVELOPMENT

The Warranted Rate of Growth

The basic question then is whether there is in fact some "warranted" rate of growth which is necessary to maintain full employment. The possibility arises because investment has a twofold function in the economic

system. It determines the immediate level of output and income on the assumption of a stable consumption function. But it also contributes to the growth of *capacity,* and hence necessitates still more investment in the next period if full employment or capacity output is to be maintained. The point is illustrated in Fig. 26, which reproduces the principle of

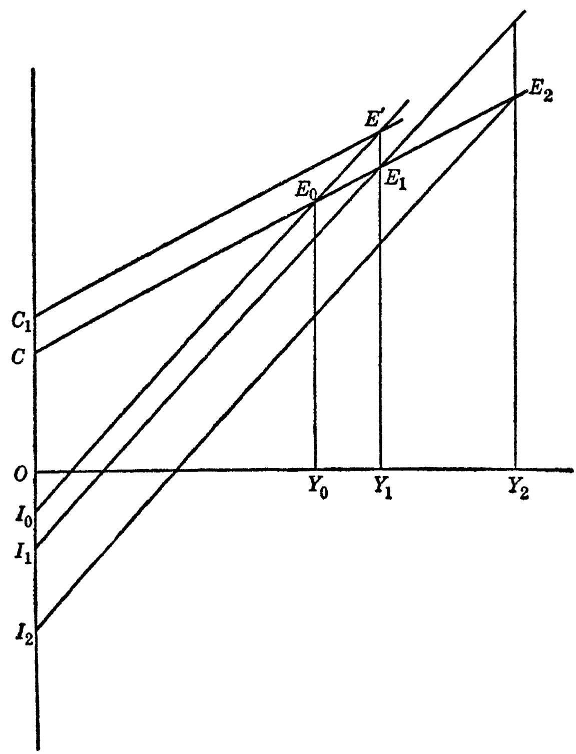


Fig. 26. The Warranted Rate of Growth

Fig. 24. Suppose that we begin with an investment, OI_0, and equilibrium output, OY_0. Now suppose, however, that as a result of this investment, the capacity of the system rises, the new capacity output being OY_1. In order to sustain this level of output, investment must be increased to OI_1. This amount of investment, however, increases capacity output still further, to OY_2, and in order to sustain this output, investment must now be increased to OI_2. If full employment is to be sustained, then, it is clear that both investment and output must increase from year to year. It is impossible to maintain full employment at a stationary level as long as there is any investment at all. The increase in both investment and output

which is necessary is geometric rather than arithmetic, so that the absolute increase in each one increases every year.[4]

Growth Cycles

In this system it is clear that there is only one pattern of growth, or growth curve, of both investment and income which will maintain full employment. Suppose now that this is *not* the growth curve which the economy as a whole can maintain by reason of its rise in population and improvement in techniques. Then clearly the system is in for trouble of some sort. If the full-employment rate of growth is greater than the maximum rate of economic development, the system will fall into cycles of rapid growth followed by depression. From a position of less than full employment, output and investment can grow at the full-employment rate, which is faster than the rate of growth of capacity of the system (the rate of economic development) because unemployed resources are being drawn into the system. Growth at this rate, however, will eventually bring the system to full employment, after which it cannot continue to grow at the full-employment rate, because the rate is now limited by the rate of growth of capacity. If this rate, however, is below the full-employment rate, full employment cannot be maintained, because investment will have to fall below what is necessary to yield full employment. Hence, instead of maintaining a steady rate of growth at the economic development rate, growth will fall below this rate as unemployment develops, and may even pass into a decline, with a large volume of unused resources. When the decline has proceeded far enough, it becomes possible to grow at the full-employment rate once again, and a new revival commences.

This system is illustrated in Fig. 27. Time is measured on the horizontal and national income on the vertical axis. *DD′*, we suppose, is the "economic development curve" which shows the growth of capacity of the system. Suppose now we start a point P_0 and grow at the full-employment rate following the line P_0P_1. In order to maintain full employment at P_1, growth would have to continue along the line P_1P_2. This, however, is impossible because of the limits of the capacity of the system; from P_0 to P_1 the growth has been possible because of the absorption of unused resources; at P_1 full employment is reached and there are no unused re-

[4] Suppose the consumption function is $C = C' + cY$, and suppose that a unit of investment leads to an increase in capacity of s. If Y_0, Y_1 are capacity outputs for years 0 and 1, we have $Y_1 - Y_0 = sI_0$. We also have the identity $I_0 \equiv Y_0 - C_0 = Y_0(1 - c) - C'$, whence $Y_1 - Y_0 = s[Y_0(1 - c) - C']$, or $Y_1 = Y_0[s(1 - c) + 1] - C's$. The solution of this difference equation is $Y_n = h^nY_0 - [\frac{(h^n - 1)}{h - 1}]\, C's$, where $h = s(1 - C) + 1$. Similarly, it can be shown that $I_1 = hI_0$, whence $I_n = h^nI_0$. In this system, therefore, I must increase at a constant rate, and if $C' = 0$ (i.e., if consumption is proportionate to income), Y will also increase at a constant rate.

sources. Further growth can only proceed along the line P_1D'. This rate of growth, however, does not demand enough investment to yield full employment; consequently it is not stable, unemployment will develop, and income will fall back along the line P_1P_3. At some point P_3 it will become possible to resume the rapid rate of full-employment growth, and income will rise till it hits the economic development curve again at P_4, whence it rebounds to P_5, and so the cycle goes on.

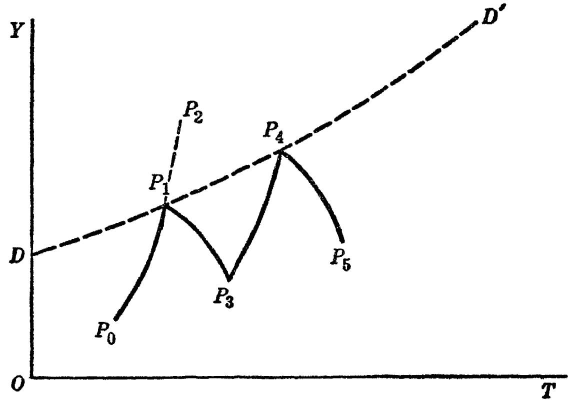


Fig. 27. Growth Cycles

Stagnation in Underdeveloped Societies

This in essence is the theory of economic instability put forward by Harrod, Hicks, and Domar.[5] It is an interesting and in many ways a useful model. It has interpretative value in explaining not only the tendency of a developed society to unemployment but also the tendency of an undeveloped society to stagnation and inflation. Thus, suppose that the consumption function is so high that the full-employment rate of growth is *less* than the rate which would be possible through economic development. The society will reach full employment at a level of investment which is less than that which is technically possible. Any attempt to force the growth of the society beyond the full employment level will result in inflation, this being the response of the system to overabsorption.

Weaknesses of the Model

The assumptions of this model, however, must be kept clearly in mind, as they may not be fulfilled in reality, and may lead to quite erroneous

[5] R. F. Harrod, *Towards a Dynamic Economics,* New York, Macmillan, 1948.
J. R. Hicks, *The Trade Cycle,* New York, Oxford University Press, 1950.
E. Domar, "Capital Expansion, Rate of Growth and Employment," *Econometrica,* April, 1946.

conclusions. The basic assumption of this model is that of so many Keynesian models—stability in the consumption function. We do not have to assume, as Harrod and Domar do, that the consumption function is a straight line through the origin (consumption proportional to income). This assumption simplifies the mathematics, but the conclusions are not destroyed if we assume a more general form for the consumption function, as we have done above. If, however, the consumption function is *not* stable, and especially if it exhibits an upward drift in the course of time, the conclusion of the model that investment must grow at a constant rate to maintain full employment is no longer valid. Suppose, for instance, in Fig. 26, that in year 1 the consumption function rises from CE_1 to C_1E'. The full-employment equilibrium will be at E', with a volume of investment still equal to that of the year before, OI_0. Depending on the rate of rise of the consumption function, then, full employment can be maintained at varying rates of growth of investment, even at zero or negative rates.

Upward Drift in the Consumption Function

The excessively gloomy predictions of the Harrod-Domar model, therefore, are not necessarily true, although any optimism depends on the empirical question: Does the consumption function, in fact, have an upward drift? There is some evidence that it has, at least in the United States, and that the long period of prosperity following the end of World War II is in large measure to be explained by the upward movement of the consumption function. This movement, in turn, may be accounted for by a number of factors, and it is not easy to identify the contribution of each. Unquestionably, the rise in liquid assets not offset by a proportionate increase in prices, as a result of the cash deficits of war finance, had some influence. Another important but much neglected factor is the age distribution of the population. Generally speaking, both the young and the old are low savers—they generally consume more than they produce. By contrast, those in the middle age groups who are in the labor force tend to be high savers—they produce more than they consume. A shift in the distribution of the population away from the middle groups toward both the young and the old, therefore, should have an effect in raising the level of the consumption function. This, however, is precisely what has been happening in the United States since 1930. The proportion of old people has been rising as a result of the increase in the expectation of life. The proportion of children has also been rising as a result of increased birth rates. Correspondingly, the proportion of people of working age has fallen. In 1930 the proportion of the population in the middle years—say from 15 to 65—was abnormally high; falling birth rates had

reduced the number of children, and the effect of the rising expectation of life had not yet worked itself out in rising numbers of old people. This fact alone is almost enough to account for the severity of the great depression. We have already noted (p. 130) that a somewhat similar phenomenon—a drastic fall in the proportion of population of working age—has affected the poor countries of the tropics and has raised their consumption functions to the point where development has become extremely difficult.

Keynesian Models in Economic Development

Even though the Keynesian models of this chapter throw some light on the problem of how a poor country gets started on the road to persistent growth, we must be careful not to apply them carelessly, for many of the problems of these countries bear little relation to the problem of underemployment in a rich country for which these models were originally devised. Thus, many poor countries today suffer from what might be called "non-Keynesian unemployment," especially in the cities. This is due mainly to the very rapid growth in population of these cities, sometimes because of refugees from war or from relocation of frontiers, and almost always because the rapid population increase in rural areas forces people out of these areas into the cities, simply because the rural economy is only adjusted to a fairly constant labor force. Thus, in a country where a quarter of the people live in the cities and the rate of growth of population is 3 percent per annum, if the rural population remains unchanged, the cities will grow at 12 percent.[6]

It is very hard to expand industrial and other urban employment at such a high rate, especially when a class of managers and entrepreneurs must be developed and when physical investment requirements are heavy. The problem is aggravated by the fact that a large proportion of the newcomers to the city are young people who have grown up in the country and who may not, therefore, have the skills of urban life. It is not surprising that under these circumstances structural unemployment develops which is not due to any deficiency in aggregate demand—indeed these countries frequently have an excess of aggregate demand and inflation at the same time that they have a large volume of urban or dis-

[6] Let the total population, p, = the city population p_c, + the rural population, p_r, and let $p_c = rp$. Then if k is the rate of growth of the total population and p', p'_c, p'_r are the values of p, p_c, p_r after one year, and the rural population is unchanged, so that $p'_r = p_r$, we have

$$p' = (p_c + p_r)(1 + k) = p'_c + p_r$$

whence

$$p'^c = p_c + kp = p_c + \frac{kp_c}{r} = p_c(1 + \frac{k}{r})$$

That is, $\frac{k}{r}$ is the rate of growth of the cities.

guised rural unemployment. The Keynesian remedies are quite inappropriate in these situations.

The Proper Use of Models

The moral of these considerations is that economic models, especially when it comes to the interpretation of a phenomenon as complex as the dynamic fluctuations or growth of the economy as a whole, are aids to thought and not substitutes. We must beware of too great reliance on the too simple model, for the assumptions of such models are always likely to be falsified. Nevertheless, without the aid of these models, the complexity of the subject matter leads either to complete bewilderment or to a retreat into the rituals of pure empiricism and the endless recording of data whose meaning always escapes us, or it leads to thinking with the aid of implicit models which we never openly recognize, and which therefore are all the more likely to lead us astray.

QUESTIONS AND EXERCISES

1. Some writers (e.g., W. S. Jevons) have tried to link business cycles with a 10- to 11-year cycle in sunspot activity. Later empirical work has largely discredited the theory. Nevertheless, assuming that the sunspot cycle creates a cycle in the yield of crops, how might this affect the economic system according to the various models outlined in this chapter?
2. Professor A. C. Pigou attributed industrial fluctuations to "waves of optimism and pessimism" among businessmen. Trace the possible effects of such waves on the various magnitudes of the economic system.
3. Much emphasis was formerly laid on the rate of interest as a governing factor in explaining cyclical fluctuations of output and income. What effect might a rise or fall in the rate of interest have on (a) accumulation, (b) consumption, (c) production, (d) the quantity of money, (e) velocities of circulation, (f) income?
4. "In order to explain the general business cycle it is not sufficient to explain how there can be cycles in particular industries; we must explain why these particular cycles do not offset each other." Discuss.
5. Suppose the age distribution of automobiles results in a sharp decline in demand for new automobiles in a certain year. What repercussions would this have on the rest of the system? How might the depression in the automobile industry become general? What would have to be done to prevent it from becoming general?
6. Under a gold-standard system a depression in one country spread very quickly to others. By what mechanism? How far do you think this might explain the breakdown of the gold standard? What might a country do to insulate itself from depressions originating abroad?

7. "Any increase in money incomes will raise the demand for labor. The way to cure unemployment, therefore, is to increase the quantity of money." What other things must be true before we can agree with this proposition?
8. The process of recovery from a depression is characterized by a shift to the right of all money demand curves, because of the increasing desire of people to spend. In the early stages of recovery this change in demand is reflected principally in a rise in employment and output. In the later stage (the boom) it is reflected in a rise in prices. Show by diagrams that these phenomena can be explained if we assume (a) that in the recovery period supply curves are very elastic and do not shift; (b) that in the boom period supply curves are inelastic, and tend to shift to the left because of increasing money costs.
   Are these assumptions reasonable?
9. A depression manifests itself in agriculture by low prices and stable or even high outputs, and in industry by low outputs and relatively stable prices. How would you explain this difference, and how might this phenomenon be related to various models of the business cycle?
10. Combine in a single diagram Figs. 24 and 25, measuring output to the right of the origin, consumption upward from the origin, investment downward from the origin, and profits to the left of the origin. Describe and investigate the properties of the system diagrammed.
11. It has been argued that as economic development is essentially a structural problem, involving reallocation of resources among various industries and a drastic change in the relative price structure, the Keynesian models do not throw much light on the processes by which it is achieved. Evaluate this argument.
12. Series from 1929 for the gross national product of the United States and its various components may be found in recent issues of the *Economic Report of the President* or in the *Survey of Current Business*. Graph the more important series, and interpret them in the light of the various macroeconomic models. Pay especial attention to (a) the great depression (1929–37), (b) the 1937–38 episode (c) the war economy (1939–45), (d) the postwar economy (1945——). Calculate for each year the *proportions* of the various components of the GNP, expressed as a percentage of the total GNP, for the two major breakdowns—(a) sources, and (b) distributional shares. Graph the results.

CHAPTER 10

THE IMPACT OF GOVERNMENT

GOVERNMENT AS AN ECONOMIC ORGANISM

By far the greatest single economic organization is the national state and its subsidiary organs. This is true even in capitalist society, where, for instance, the governments, Federal, state, and local, of the United States in 1964 absorbed around 21 percent of the GNP. It is impossible, therefore, to regard the economic system as something apart from political organizations, though it is useful to regard the system as consisting of two sectors, a private sector of nongovernmental organizations and a governmental sector.

The Framework of Legal Prohibitions

The impact of government upon the economy, and especially upon the private sector, is felt through two main channels. In the first place the government, in its judicial and legislative aspect, sets the legal framework within which private organisms operate. This it does mainly by means of prohibitions enforced by legal penalties—the commandments of government, like those of Moses, are mostly concerned with prescribing the things which we may not do. These prohibitions may affect economic life in very far-reaching ways. There may be, for instance, prohibitions of entire industries, as under the 18th Amendment. The production of dangerous or adulterated commodities may be prohibited, as under the Food and Drug Act. Certain types of securities may be proscribed, as under the Securities and Exchange Act. Various forms of property are defined, and violations of ownership rights, prohibited. Certain types of agreements between individuals may be prohibited, such as agreements in restraint of trade, under the Sherman Act; or yellow-dog contracts, under the Wagner Act. Certain types of economic behavior or policy are prohibited, such as price discrimination, under the Robinson-

Patman Amendment to the Clayton Act; or payments to "stand-ins" for work not performed, under the Taft-Hartley Act.

There may also be restrictions on the prices at which transactions may legally be performed. It may be illegal to buy or sell labor below a minimum wage, and it may be illegal to buy or sell commodities above a maximum ceiling price. It may be illegal to hold certain types of assets (e.g., gold). All these legal restrictions set the framework within which private organizations operate. The framework is not absolutely rigid, as indicated by the existence of bootlegging under prohibition or of black markets under price control. Illegality is an obstacle, a "cost" of performing the illegal acts, and from the point of view of its effect on the total economic system, it operates very much like "natural" obstacles.

Transactions of Government

All the activities of government in the setting up and enforcement of prohibitions of one kind or another are negative in that they determine the course of economic activity only indirectly by setting up or removing certain obstacles. There is, however, another economic aspect of government which is "positive" in the sense that it consists of the direct economic relationships which government, as an economic organism, has with the other organisms of the system. In this sense government is to be regarded merely as a very large and peculiar kind of "firm" with its own system of relationships with the other organisms of society.

Exchanges and Transfers

The transactions of government with private firms or households can be divided into two important classes—*exchanges* and *transfers*. A government makes exchanges just like any other economic organism. It purchases labor, and pays wages in return; it purchases goods of all kinds and pays money to the sellers. It also sells certain goods and services—e.g., postal services, war surplus goods—for which it receives payment in money from the buyers. In making exchanges a government differs from a private organization only in its ability to neglect to some extent the effect of its exchanges on its asset structure. Governments have two powers which private organizations do not usually possess: the power to tax and the power to create money. Because of these powers the transactions of government include an important class of *transfers*, which differ from exchanges in that they are one-way transactions. An exchange is a two-way transaction; it involves the mutual transfer of assets between the exchangers. Thus, when a government purchases goods, money passes from the government to the private seller, and goods of equal book value pass from the seller to the government. From the point of view of the seller, the exchange is an asset transformation; he has replaced, say,

$50,000 in goods on the assets side of his balance sheet by $50,000 in cash. A government transfer, on the other hand, involves simply the transfer of assets of some kind—usually cash—from the private organism to the government or from the government to the private organism. The commonest form of transfer is, of course, a tax, which is a one-way transfer of money from private individuals or organisms to the government, resulting in a net diminution of the assets of the taxpayer. There are also "negative taxes" in the form of subsidies, pensions, grants, etc. which consist simply of a money payment from the government to private persons or organizations, without any transfer in return, thus adding to the assets of the recipient.

Government and Payments System

Let us consider first the impact of government on the payments of the private economy. The "balance of payments" of government is its "cash surplus." If the cash budget is balanced, it means that government is taking in from all sources just as much money as it is paying out to all recipients, no matter whether the payments are transfers or exchanges. If there is a deficit in the cash budget, the government is paying out more money than it is taking in. The difference clearly must be equal to the increase in the cash holdings of private persons and organizations, which in turn is equal to the favorable balance of payments of the private sector of the economy. This conclusion applies strictly only in a closed society; in the case of a single nation it is still true that a deficit in the cash budget of government means an increase in holdings of that government's money, but part of this increase may be held by foreigners. Thus a deficit in the government's budget must result in a *surplus* in all other cash budgets taken together. Similarly, if the government has a surplus in its cash budget, it is taking in more than it is giving out, and therefore the private sector will find its stocks of cash diminishing and will find itself with net budget deficits. These principles can be perceived quite clearly from the payments tables such as Table 4 in Chapter 4. If one of the individuals—say, A—is supposed to be the government, it is clear that A's budget deficit (excess of expenditures over receipts) must be equal to the net sum of all the other budget surpluses (excesses of receipts over expenditures) according to the principle that the sum of all receipts must be equal to the sum of all expenditures.

Government's Power to Create Money

Government differs from private institutions, from the point of view of the payments system, mainly because of its much greater capacity for the creation of money or, what is almost the same thing, of readily

marketable securities. It should be observed that private organizations possess this power to a limited extent. Thus, a firm that wishes to undertake an investment, in the course of which it will be necessary for a time to have a deficit in its cash budget as it pays out more than it is receiving, may finance this deficit by the creation and sale of its securities, either to the public or to a bank. Governments likewise may create and sell securities, either to private individuals and corporations or to a bank. The effects are very similar in both cases. Where securities are sold to private individuals or to nonbanking corporations, there is a corresponding transfer of money to the body which sells the security. Thus, there is a new asset in circulation in the form of the security, but no new money has been created. When, however, a government or any organization sells its newly created securities to a bank, the bank purchases these securities with newly-created bank deposits. Unless its reserve position forces the bank to liquidate some other assets, there is net creation of money in the form of bank deposits as well as creation of securities. This is the usual method of financing a governmental deficit in these days, rather than by the creation of government money; the government creates government bonds, which it sells to the banks for deposits to be paid out to private persons and firms. The effect, however, is much the same as if the government had created money directly, except that the banking system increases its holdings of government bonds and its total deposits. The difference between a government and a private corporation lies in the taxing power of government. As long as the people have faith in the government's taxing power, its capacity to issue and sell securities is almost unlimited, as its power to create legal tender is also unlimited. The power of private firms to issue securities depends on the belief of potential purchasers in the profitability of the enterprise. But whereas the profitabilty of private enterprises depends on their ability to wheedle money out of their customers by offering services, the profitability of government has behind it the big stick of the taxing power.

How Government Can Regulate the Payments Total

We can see now how the fiscal system of government could be used to regulate the total of private payments. If private individuals on balance are trying to accumulate money, the government can prevent the decline in payments which such an attempt involves by having an appropriate cash deficit and so increasing the quantity of money in the hands of the public. If private individuals are trying to decumulate money, government can prevent the rise in the payments total which would result by having a cash surplus and so absorbing money from the public. It should be perfectly possible, therefore, given adequate in-

formation and a flexible payments system, for government to stabilize the total payments within as narrow limits as it wishes. The simplest way to do this would be through the tax system; if tax levies were automatically reduced whenever the total of private payments fell, and automatically raised whenever the total payments rose, a "governor" or cybernetic mechanism would be introduced into the payments system capable of stabilizing it within reasonably narrow limits.

Automatic Tax Adjustment Through Progressive Taxation

If the tax system is progressive, in the sense that people with higher money income pay a larger proportion of this income in taxes, the tax system then becomes an "automatic stabiliser." Incomes are closely, though not exactly, related to the total volume of payments. Hence, if payments shrink, incomes decline, and if the tax system is progressive, tax payments will decline faster than incomes, for at low incomes people pay a smaller proportion in taxes. Therefore if government expenditures do not decline, there will be a government cash deficit, which will increase the stock of money and so increase the volume of payments. Similarly, if there is inflation and payments increase, incomes will increase, and tax collections will increase in larger proportion, thus producing a cash surplus and a withdrawal of money from circulation. This stabilizing property of the tax system only operates if taxes are collected concurrently with income earned. Thus if, for instance, income tax this year is paid on last year's income, then the tax system can easily become destabilizing, for when income is falling, tax collections, being based on the larger income of last year, may actually rise as a proportion of income. Perhaps one of the most important social inventions of the twentieth century has been the "deductable at source" and currently payable income tax, which, combined with the progressive rate of tax, has made the income tax an important automatic stabilizer in the United States and other developed countries.

It is a moot question whether the present tax system is a sufficiently powerful stabilizer, and whether additional stabilizing features should be added, such as, for instance, an automatic cut in tax rates when income declines, or a rise in rates when income rises, above, say, a certain threshhold rate. The use of a tax cut in 1963 to stimulate a rise in income suggests that we are beginning to overcome the psychological and political obstacles which lie in the way of the use of the tax system as a deliberate stabilizer. We still, however, seem to be some way from accepting the idea that the tax system can be used automatically for this purpose, and, indeed, the principle of close supervision of the tax system by the legislature still has much to recommend it when we come to the

details and the incidence of taxes. The problem here is how to leave to the legislature the task of determining the incidence and kinds of taxes while at the same time linking the total *amount* of taxes collected, which is the most crucial variable in stabilization policy, to the stabilization needs of the economy at large. One possible way of doing this would be to have a special kind of money ("yellow dollars") for paying taxes, and then set up a rate of exchange between ordinary ("green") money and yellow money which would vary, perhaps even automatically, with the stabilization needs, the price of yellow money rising in time of inflation—so that more green money would be needed to pay tax obligations—and falling correspondingly in times of deflation.

GOVERNMENT AND EMPLOYMENT

Government Absorption

The effect of government on the total of employment and output is more complicated than its effect on payments. Nevertheless, certain broad principles can be enunciated. In the first place, government is highly significant as an "absorber" of output. The net purchases by government (ie., purchases minus sales) constitute governmental absorption of output. This absorption, like that of households and businesses, is part consumed and part accumulated. Government purchases, it should be noted, include the labor of civil servants, the product of which, for the most part, can be regarded as government consumption. From the point of view of employment and output, however, the distinction between governmental consumption and accumulation is unimportant except in the case of material, such as surplus war material, which is accumulated and sold back to the public, or in the case of government investment, for example, in dams and power plants, which competes with and hence may discourage private investment.

Its Effect on Employment

If we can assume simple functions relating total output and private consumption and accumulation, it is not difficult to calculate that amount of governmental absorption which will result in full employment. If, at full-employment output and incomes, the amount of private consumption is C and of private accumulation is A, then the amount of governmental absorption necessary to secure full employment, G, is $G = P - C - A$, where P is capacity (full employment) output. Unfortunately, of course, the problem is by no means as simple as this formula would indicate; the consumption-output and accumulation-

output functions are not stable enough to be discovered statistically with any degree of confidence, and are also affected by other operations of government—e.g., its balance of payments—and by innumerable obscure psychological and sociological conditions. Thus an increase in governmental absorption designed to create full employment may have such an effect in restricting private accumulation and consumption, if the general sociological atmosphere is unfavorable to governmental expansion, that the net effect of even an expansion of governmental absorption is unfavorable to output. On the other hand, where the expansion in governmental absorption—e.g., in a popular war—is approved by the general community, it may have favorable effects on private accumulation and consumption and may rapidly produce conditions of "pressure economy."

Graphic Solution. The graphic solution to the simple system is shown in Fig. 28, where output is measured along *OK,* private absorption

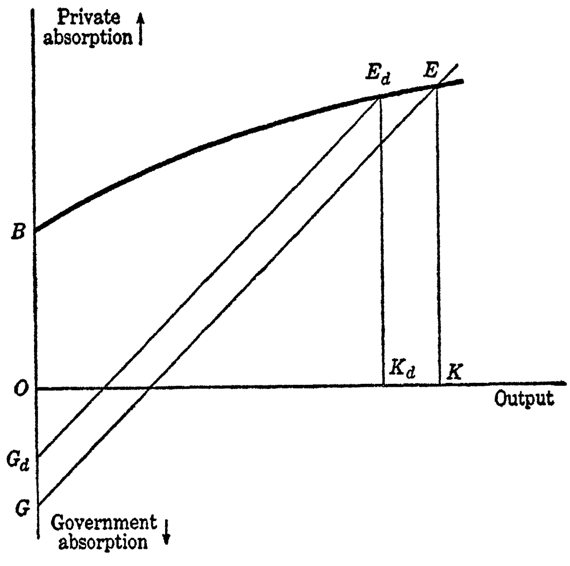


Fig. 28. Output Equilibrium with Government Absorption (1)

along *OB,* and government absorption along *OG. BE* is the private absorption curve, showing how much will be consumed and willingly accumulated by the private sector of the economy at each level of output. Suppose now that *OK* is the capacity, or full-employment output. *KE,* then, is the amount of private absorption at full employment. We now draw a 45 degree line from *E* to meet the vertical axis in *G. OG* is,

then, the amount of government absorption required to yield full employment, as only if government absorption is OG will the whole product be absorbed $(KE + OG = OK)$.[1] If government absorption is less than OG, say OG_d, the equilibrium output will be at less than capacity, OK_d. From a position of unemployment there will be a "government multiplier," m_g, exactly analogous to the investment multiplier of page 57, defined as the increase in the equilibrium output which results from a unit increase in government absorption. If α is the private propensity to absorb (the slope of the curve BE), then as before we have

$$m_g = \frac{1}{1 - \alpha}$$

The multiplier is again a measure of the sensitivity of the system; if it is large, quite small changes in government absorption may produce large changes in equilibrium output.

If government absorption is so large that the equilibrium output in the basic model is above capacity, there will be inflation, as the economy attempts to consume more than is available for consumption. As we have seen earlier, the inflation will only be successful (i.e., self-limiting) if it succeeds in lowering the private absorption curve to the point where there is no longer unwanted decumulation.

Model with Disposable Income and a Government Deficit

The simple model outlined above has one grave weakness: it postulates private absorption, B, as a function of *total* income, whereas it may be more realistic, especially as far as the consumption component is concerned, to suppose that private absorption is a function of private, or disposable, income, Y_d,[2] which is total income, Y, minus total tax collections, T. We then have a model with the following equations, G, being government absorption:

1. The basic identity, $Y = B + G$
2. The private absorption function, $B = F_b(Y_d)$
3. The definition of disposable income, $Y_d = Y - T$

If G and T are given, these three equations suffice to determine the three unknowns, Y, Y_d, and B.

[1] In algebraic terms the model has three unknowns: government absorption, G., private absorption, B, and national output, Y. The three equations are

1. The basic identity, $Y \equiv B + G$
2. The private absorption function, $B = F_b(Y)$
3. Government absorption given, $G = G$

[2] This corresponds roughly to the gross disposable income concept, A'_7, on p. 34.

The system is illustrated graphically in Fig. 29. Disposable income is measured along the horizontal axis from *O*, private absorption is measured upward from *O*, government absorption and taxation downward. *CE* is the private absorption curve. Then let *OG* be total government absorption, *GF* be the total tax bill. *OF* is then the government deficit, for government absorption is what is bought with government expendi-

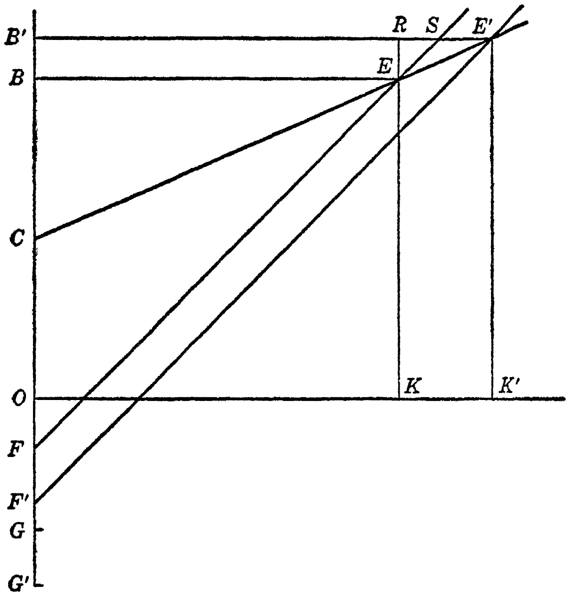


Fig. 29. Output Equilibrium with Government Absorption (2)

ture, and is therefore measured in dollar terms by government expenditure. If now we draw a 45 degree line from *OF* to cut *CE* in *E*, *E* is the point of equilibrium, *OK* is the equilibrium disposable income Y_d, and if a perpendicular *EB* is dropped to the vertical axis, *GB* is the total income, *Y*. From the geometry of the figure it will be seen that all the equations above are satisfied: (1) $GB = OB + GO$, (2) *E* is on the private absorption curve, $B = F_b(Y_d)$, and (3) $OK = EK + FO = FB = GB - GF$.

Effect of Pure Increase in Deficit

Now let us suppose that there is an increase in the deficit from *OF* to *OF′* without any change in government absorption. That is, tax collections fall from *GF* to *GF′*. Drawing the 45 degree line from *F′* to cut the private absorption curve in *E′*, we find the new position of equilibrium, *E′*. Private disposable income has risen to *OK′*, total income

to GB'. It is not difficult to derive "deficit multipliers" in this case. If m_d is the disposable income deficit multiplier and α is the propensity to absorb, we have

$$dY_d = KK' = RE' = RS + SE' = ER + FF' = \alpha dY_d + dF,$$

whence

$$m_d = \frac{dY_d}{dF} = \frac{1}{1 - \alpha}$$

If now m_t is the total income deficit multiplier, we have

$$m_t = \frac{dY}{dF} = \frac{BB'}{dF} = \frac{RE}{dF} = \frac{\alpha dY_d}{dF} = \alpha m_d = \frac{\alpha}{1 - \alpha}$$

We recall that α must be less than one if the system is to be stable, so that the total income multiplier will be less than the disposable income multiplier.

Effect of Increase in Deficit, Taxes Unchanged

Consider now the case where there is an increase in the deficit, not as a result of a reduction in taxes, but as a result of an increase in government expenditure or absorption. The deficit moves as before from OF to OF', but now government absorption also rises by an equal amount from OG to OG', taxes $(G'F')$ being the same as before (GF), so that $FF = GG'$. The change in disposable income is the same as before, being governed in this model only by the deficit, and not by the volume of government absorption. The change in total income, however, is now equal to the change in disposable income, as taxes are unchanged. The change in total income is $G'B' - GB = BB' + G'G = ER + F'F = RS + SE' = RE' = KK'$, the change in disposable income. In this case, then, the total income multiplier is the same as the disposable income multiplier.

Effect of Rise in Government Absorption, Deficit Unchanged: The Ricardian Case

Suppose now that we have a rise in government absorption, financed by an increase in taxes so that the deficit remains unchanged (it makes no difference to the analysis, of course, if we assume the deficit to be zero). Now we ask ourselves what must happen to the private absorption function if the equilibrium output is to remain unchanged. This might be described as the "Ricardian" case, as Ricardo always assumed in effect, that a rise in government absorption would not raise employment, as private absorption would fall by an equal amount. This is illustrated in Fig. 30, which is essentially the same diagram as Fig. 29. Here we suppose

an increase in government absorption from OG to OG_1, and of tax collections from GF to G_1F, so the deficit remains unchanged at OF. Now, however, we want total income to remain unchanged. Private absorption, therefore, must fall from OB to OB_1, where $BB_1 = GG_1$, in order that the new total income, G_1B_1, shall be equal to the old, GB. This, however, in the absence of other changes requires a fall in the private ab-

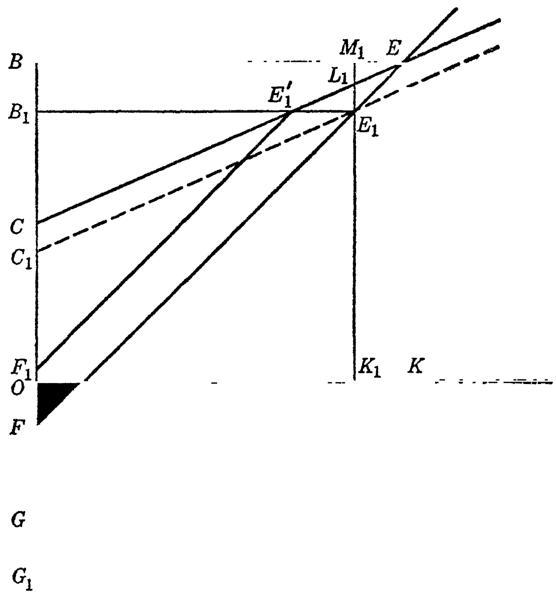


Fig. 30. Output Equilibrium with Government Absorption (3)

sorption curve to go through the point E_1, where the horizontal line from B_1 cuts the 45 degree line FE. C_1E_1 is one such new private absorption curve, though any other reasonable curve which goes through E_1 would be satisfactory. It will be observed that the decline in private absorption at the disposable income, OK_1, L_1E_1, is less than the rise in government absorption, GG_1. $L_1E_1 = E_1M_1 - L_1M_1 = B_1B - \alpha M_1E = GG_1 \ (1 - \alpha)$.

The Principle of Overfinance

Now suppose that the private absorption curve does *not* fall and we ask ourselves what we have to do in order to prevent total income from rising under the impact of expanded government absorption. The answer is clearly that we must diminish the deficit by increasing tax collections. In Fig. 30, B_1E_1 cuts the private absorption curve, CE,

in E'_1, and we draw a 45 degree line from E'_1 to cut the vertical axis in F_1. OF_1 is the deficit (in the case of the figure, a surplus) which is necessary to prevent a rise in total income when government absorption rises from OG to OG_1. That is to say, it is not sufficient to raise merely enough in taxes to cover the increased government expenditure. The rise in tax receipts must be *greater* than the rise in expenditure if total income is not to increase. This may be called the "principle of overfinance." It is, of course, a corollary of the principle already enunciated that expansion of government absorption which is financed wholly by increased taxes is still inflationary and will produce an increase of total income. Starting from a position of underemployment, of course, this will probably mean increased real output; from a position of full employment, however, increased government expenditures must be overfinanced if they are not to give rise to inflation. It is not surprising in the light of this principle that war nearly always gives rise to inflation, for war requires generally a sharp increase in government absorption. It requires considerable financial virtue not to increase the deficit at such a time, and the amount of financial rectitude required to diminish the deficit, and perhaps to run a surplus, at a time when expenditures are increasing sharply is probably beyond the capacity of any known government.

Application to Effects of Private Investment

The above analysis can also be applied to the study of the effects of private investment as well as of government absorption. In so far as investment is financed either by the creation of new money through the banking system, or by a shift in money balances from business toward households, business in general will have a "cash deficit" analogous to a government deficit. If now household absorption is regarded as a function of *net* household income, that is, gross household income plus increases or minus decreases in the money stock of households, a cash deficit of businesses has the same effect in increasing the money stock of households as a cash deficit of government, for it means that business takes from households less than it pays back to them. Under these circumstances the inflationary effect of investment is augmented. The principle of overfinance also applies to investment. If there is an increase in investment there must be an actual decline in the money stocks of households if equilibrium income is to remain constant. This perhaps is the principal explanation of the frequently observed phenomenon—that full employment tends to pass into a price-wage inflation. If investment increases *from* a position of full employment it is extremely unlikely that the decrease in consumption will just offset the increase

in investment. In order for this to happen there would have to be a sharp reduction in household money stocks—something which the existing financial and fiscal system is most unlikely to provide!

Assumptions of the Model

In all these models the warning should again be issued that the validity of the conclusions depends in large measure on the degree to which the behavior functions are in fact stable over time, and the possibility of shifts in these functions must always be kept in mind. Thus the private absorption function assumed in the above analysis may shift for a variety of reasons, many of which may be connected with various aspects of government policy. In the above analysis, for instance, we have been concerned only with the total volume of tax payments, not with the method of collecting taxes, and it will obviously make a difference to the private absorption function whether taxes are levied on income, on sales, on households, on corporations, or on imports. Nevertheless the above models provide a rough outline for the understanding of the overall aspects of public finance, and a method of analysis of the secondary effects through their impact on the primary behavior functions.

WAR FINANCE AND INFLATION

The great rise in government absorption during World War II, at least part of which seems to be permanent, involved the world in a period in which inflation rather than the unemployment of the over-thrifty 1930s has been a dominant problem. Inflation has been mainly the result of a volume of government absorption more than sufficient to yield full employment, and inadequately offset by tax collections. The price inflation which resulted is itself a form of taxation, though often a very arbitrary and unjust form. In a "real" sense taxation is the decline in private absorption which must take place if government expands its absorption from a position of full employment. This decline can be accomplished *either* through the money tax system *or* through inflation. It can be accomplished through the tax system by taking enough money away from the private sector to reduce private absorption to what is "available" after government has absorbed the amount of product it wants. Or, if taxes are not high enough to accomplish their proper task, inflation will, as we have seen, redistribute income away from the low savers toward the high savers and so will lower the total private absorption function (see p. 65).

Open and Suppressed Inflation

Inflationary pressure can take two forms—"open" or "suppressed." In an open inflation there is no control of prices, and they will rise until the redistribution of income is sufficient to accomplish the necessary reduction in private absorption. In suppressed inflation, prices and money wages are held down by law, and, as a result, shortages develop—the "pipelines" empty, and consumption is restricted simply because people cannot find things to buy. These shortages can be regularized through rationing, in which consumption is restricted not by high prices but by direct allocation to consumers by means of ration coupons or by a system of priorities.

Whether open or suppressed inflation is preferable seems to depend on the degree of the inflationary pressure. Price-wage control and rationing in themselves have a high social cost in terms not only of resources devoted to administration and evasion, but also in the inherent difficulty of administering the system without increasing maladjustments in the *relative* price structure. Unfortunately we cannot control the general level of prices and wages without controlling particular prices and wages—the general level is a statistical abstraction, and law can only operate on the concrete reality of particular prices and wages. The method of price control, therefore, is always to "freeze" the price structure as it existed at some date, and then set up a machinery for making adjustments and exceptions as the course of time and change in techniques and in demand make the relative price structure of the base date obsolete. The administrative machinery of adjustment, however, is slow and clumsy compared with the swift sensitivity of the free market, and, as a result, the longer price control operates, the more distorted becomes the relative price structure and the stronger the temptation to evade the law in black markets. On the other hand, great as are the difficulties of price control and rationing, they are almost certainly less than the evils of *rapid* open inflation. Inflation, as we have seen, is a very poor equilibrator, because, though it probably restricts private consumption, the distribution of income which it brings about increases profits and so probably stimulates investment. The net effect of inflation on private absorption (consumption plus investment) may therefore be small, or even perverse. If this is the case, inflation may be quite unable to accomplish the task of restricting private absorption and may degenerate into "hyperinflation."

Hyperinflation

Hyperinflation is the condition where all the equilibrating effect of inflation is lost, and the value of the monetary unit sinks almost to

zero. Such hyperinflations occurred in France after the Revolution, in Germany and Russia after World War I, and in Hungary after World War II. When prices are rising very rapidly, it becomes apparent even to the least speculative that it is folly to hold money for any length of time. People therefore seek to get rid of their money as soon as it is received, for even the wage earner realizes that by the end of a week his money may be worth only half of what it is at the beginning. Hence there is an enormous increase in the velocity of circulation, which may even reach the physical limits set by the speed with which people can run from the teller's window to the shop. The increase in the velocity of circulation, then, reinforces the effect of the increase in the quantity of money, and money incomes and prices skyrocket to figures which may be billions of times their former value. The owners of bonds, mortgages, government debts, pensions, and annuities then lose the whole value of their property in so far as it consists only of a right to receive certain fixed sums of money, for these sums now become worthless.

The Lag in Government Inflation

In a mild government inflation, however, it is usually observed that the velocity of circulation falls a little. Some of the new money goes into hoards and consequently does not affect prices. But once the actual creation of money ceases, the velocity of circulation frequently increases, for the money which has been put away now comes out and is spent. So we often find that the creation of money during the war does not affect prices as much as it might, but that when the war is over, even if the creation of new money ceases, the velocity of circulation increases as the hoards are released, and prices undergo a further sharp rise. This phenomenon was observed after the Napoleonic wars, and also after World Wars I and II. It goes far to explain the postwar boom and the subsequent depression.

Forced Saving

The phenomenon of "taxation through inflation" is a special case of a more general phenomenon. Any rise in prices, whatever its cause, "taxes" the people whose money incomes have not risen in proportion in that it reduces their real income. This reduction in real income through a rise in prices is sometimes called "forced saving," though the term is not a good one owing to the ambiguity in the word "saving." It may be caused not only by government inflation, but also by a rise in the volume of bank credit, or even by a spontaneous rise in the velocity of circulation. Thus we have seen how an "investment inflation" is something like a government inflation. In this case also

there is forced saving in the sense that there is a restriction of consumption through the redistribution of real income which the inflation produces. There is a real problem of economic justice here, because when investment is financed by inflation the people who come into possession of the new assets are *not* those whose enforced parsimony made the investment possible. The overall increase of real assets is always a result of saving—i.e., consuming less than is being produced—but the people who have restricted their consumption are not necessarily the ones that own the new assets.

GOVERNMENT DEBT

Government debt usually arises when government receipts from taxation and sales of goods and services fall below government expenditures. This gap may be filled in three ways—by the issue of government money (e.g., greenbacks), by the sale of government bonds to banks, and by the sale of government bonds to the general public. Government money may be regarded as a variety of national debt—i.e., as a noninterest-bearing government security. Usually, however, it is not so regarded, and the term "national debt" is reserved for interest-bearing government bonds. The sale of government bonds to banks, as we have seen, is virtually equivalent to the creation of money, for the bonds are exchanged for newly created bank deposits. The sale of bonds to the general public is in its immediate effects equivalent to taxation as far as the payments effects are concerned, for it results in a transfer of liquid funds (cash or bank deposits) from the general public to the government. The sale of government bonds differs from taxation in that the general public does not suffer a diminution in the value of its assets, but merely exchanges one form (money) for another form (government bonds). Where government bonds are redeemable at short notice, they are themselves so liquid an asset as to make this change not very significant. Nevertheless, the purpose of the exchange is to diminish the liquidity of the assets of the general public, and so perhaps relieve inflationary pressure. An individual is somewhat less likely to increase his expenditures if he has, say, $500 in cash and $1500 in government bonds, instead of $2000 in cash.

Interest on Government Debt

Government securities usually have to bear interest if people are to be induced to hold them, and this in itself creates something of a problem. The payment of interest on a government bond is an ex-

penditure—i.e., a cash outlay of government—and must itself be met by one of the ways suggested above: by taxation, by printing money, or by selling more securities to banks or to the general public. If the government balances its budget, the interest on the debt represents a transfer of money from taxpayers to the debt holders. If everyone who paid taxes owned government debt in proportion to the taxes which he paid, this transfer would mean little, for what each person disbursed in taxes out of one pocket would flow into another of his pockets in the form of interest on his government bonds. In that case the distribution of income would not be directly affected by the amount of the debt. Whether the debt were $1 million or $1 billion would make little difference to the structure of incomes. Where, however, the people who pay taxes and the people who own the debt are not the same group, government debt is an instrument for the transfer of income from the taxpayers to the debt holders. The debt holders generally belong to the wealthy and to the middle class. This remains true even when the debt is held largely by banks and insurance companies. In that case the depositors of banks and the policyholders of insurance companies are indirect beneficiaries of the interest payments on the debt. But bank deposits and insurance policies are not held to any great extent by the poorest section of the population, whereas indirect taxation and sales taxes reach down into the lowest income groups. The payment of interest on government debt, therefore, may result in a transfer of income from the very poor to the middle and upper classes—a result which is contrary to most other government policies. This transfer could be prevented by a tax system planned with that end in view, but it is doubtful whether such an objective could be reconciled with other ends of the tax system. Hence, there is strong likelihood that a large government debt will introduce a "regressive" feature into public finance.

The Conversion of Debt

Sometimes a government may reduce the interest payments which it has to make on its debt by the process known as *conversion*. This consists in borrowing at a low rate of interest in order to pay off a debt on which a high rate is paid. Suppose that at a time when the rate of interest was high, the government borrowed $100 million at a rate of interest of 5 percent per annum, the principal to be repaid (or redeemed) in 20 years. Suppose that at the end of the 20 years the rate of interest was only 3 percent per annum. Then the government could borrow another $100 million at 3 percent and pay off the old debt.

There would still be $100 million of debt outstanding, but the government would now only have to pay out $3 million in interest in each year instead of $5 million.

Advantages of a National Debt

Although a national debt may have certain adverse effects on income distribution, it may have a balance of advantage up to surprisingly large levels. We cannot, perhaps, agree with Daniel Webster unqualifiedly that "a national debt is a national blessing." Nevertheless, there is much to be said for it. It certainly contributes to political stability—a widely distributed national debt is perhaps the best insurance policy a government can have against the wrath of its people. It may also make positive contributions to economic welfare. In a period when people as a whole wish to "save"—i.e., increase their total assets—but when the increase in real assets which is economically advantageous is not great enough to supply the demand for "savings," an increase in government debt is a useful way to achieve a rise in the total of net worths without accumulation of goods or inflation of prices. In a sense this rise is fictional, for it only arises because government as such has no balance sheet and because noncurrent tax liabilities are not counted as liabilities in private balance sheets. Because of this an increase in government debt results not only in an increase of total assets but also in an increase in the net worths of the public. The issuance of private debt increases assets and liabilities together, and hence does not directly add to net worths. Because the desire to increase net worths in the absence of opportunities for accumulating real capital is one of the principal sources of unemployment, an increase in government debt may result in an increase in output.

Effect on Composition of Assets

Although the increase of the national debt increases the total of net worths, it also has an effect on the *composition* of assets, the effect of which is not easy to assess. The existence of a large volume of safe, liquid assets in the form of government bonds, even though the rate of interest on them be low, may discourage the accumulation of real capital and hence may contribute to the very malady which the growth of the national debt is supposed to cure. On the other hand, the development of a class of rentiers living off the interest on the national debt by increasing consumption and diminishing the labor force will have effect in diminishing unemployment as such, though a solution to the unemployment problem which involves the development of a

class of respectably but permanently unemployed does not perhaps recommend itself too highly.

Composition of the United States National Debt

Many people, arguing falsely from the analogy of private debt, conceive the national debt as a great burden which we are bequeathing to our children. Even in quantitative terms, however, the debt is often much overestimated. Table 24 shows the estimated ownership of the United States national debt at selected dates (from Table B56 of the *Economic Report of the President, January, 1965*). It will be observed that as a proportion of GNP, the net debt is smaller in 1963 than in 1939, and the proportion held by individuals is also about the same (21 percent).

TABLE 24. THE UNITED STATES GOVERNMENT OBLIGATIONS: ESTIMATED OWNERSHIP (In Billions of Dollars)

1. Year ended	1939	1945	1963
2. Gross public debt	47.6	278.7	310.1
3. Held by United States government investment accounts	6.5	27.0	58.0
4. Held by Federal Reserve banks	2.5	24.3	33.6
5. Held by commercial banks	15.9	90.8	64.3
6. Held by mutual savings banks and insurance companies	9.4	37.4	17.1
7. Held by other Corporations	2.2	22.2	20.6
8. Held by state and local governments	0.4	6.5	21.1
9. Miscellaneous holdings	0.7	9.1	29.2
10. Individual holdings	10.1	64.1	66.2
11. GNP	91.1	213.6	583.9
12. Net public debt[a]	38.6	227.4	218.5
13. Net public debt as percent of GNP	42.4	106.5	37.4
14. Debt held by commercial banks as percent of GNP	17.5	42.5	11.0

[a] Public debt held outside the Federal Government and the Federal Reserve Banks.

Gross public debt as a percentage of GNP was 15.6 in 1929, 35.6 in 1932, 52.3 in 1939, 130.4 in 1945, 53.1 in 1963, and 51.2 in 1964 (preliminary estimate). These figures reveal the importance of depression in increasing the burden of debt, the relative unimportance of the New Deal, the overwhelming importance of World War II in increasing the debt between 1940 and 1945, and the relative decline of the debt since 1945. Another important feature is shown in line 14, where we express the debt held by commercial banks as a percent of GNP. We see that this was very large in 1945, right after the war. This fact created real difficulties for monetary management, as the Federal Reserve Banks felt obliged to hold up the price of government securities for fear of

a collapse of the banking system if the value of government bonds were to fall sharply. Today, however, the problem has largely disappeared because of the rise in GNP against a fairly stationary (since 1947) holding of public debt by commercial banks.

The Burden of Interest

Even though the size of the national debt of the United States is no cause for alarm, and, indeed, a sharp rise in the next few years may well be in order if full employment is to be maintained, the rising burden of interest payments on both private and public debt is a matter of some concern. Interest rates, on short-term bank loans, for example, have more than doubled between the mid-1940s and the mid-1960s, and the proportion of GNP going to interest payments has also risen sharply, from 1.5 percent in 1945 to 4.1 percent in 1963. This rise is, in large measure, the result of deliberate government policy, which has stressed the correction both of domestic inflationary movements and of adverse balances of payments by high interest rates rather than by higher taxes, and the correction of unemployment by lower taxes rather than by lower interest rates. This policy is open to severe criticism, on the grounds, first, that interest is a burden on the productive parts of the society which should be kept to the minimum necessary to maintain the smooth operations of the financial system, and, secondly, that using interest rates as an attempted stabilizer penalizes investment and slows down the rate of economic growth. A combination of low interest rates and high taxes in periods of inflation, and low interest rates and low taxes in periods of deflation, will encourage investment, use consumption as the main stabilizing factor, and result in a higher rate of growth than the opposite policy of attempting to deal with inflation by high interest rates and cutting down investment.

The Economics of Disarmament

One of the most important economic problems facing the United States is the role of the military establishment in the economy, and especially the possible economic consequences of a sharp decline in the "war industry" (that part of the economy which produces what is paid for by the military budget), as a result of the growth of world order. There is a widespread myth in the United States that only the increase in the war industry got us out of the great depression and maintains us in prosperity. Unfortunately, the myth has some evidence to support it. It is true that from 1940 to 1945, the enormous rise in the war industry was achieved largely by absorbing unemployed resources. It is also true that since 1945 the United States has devoted a much larger proportion of its GNP to

the war industry than it did before the war (9.7 percent in 1963 compared to 1.4 percent in 1939). It is quite untrue, however, that this war industry is necessary to maintain full employment. In fact, it represents a substantial bite out of personal consumption, which was 74 percent of GNP in 1939, and is only 64 percent in 1963. In principle, then, the problem of how to get full employment without the war industry is very simple; if we take, say, $50 billion of war industry out of the GNP, then if we do not want a deflation, we must expand the other components by $50 billion. It should be easy to expand personal consumption by tax reduction—indeed, if we could expand personal consumption to the 1939 proportion of GNP, we would have actual inflationary pressure under disarmament! It should be easy also to expand investment, through reduction of interest rates and an "easy money" policy. It should be easy to expand civilian government activities, which are being severely squeezed and starved by the war industry. It should also be easy to expand foreign investment, public or private.

One conclusion, however, which is suggested by the principle of overfinance is that at least in the short run a sharp decline in the war industry should be offset by a larger decline in tax collections. It is hard to persuade the American public that a reduction in the government expenditure should be accompanied in the first instance by a budget deficit; once we realize this, however, the problem of disarmament loses most of its economic terrors. There are, of course, real problems of adjustment—that is, the actual transfer of industries from war to civilian output—and this might be severe in particular localities. The remarkable experience of the great disarmament of 1945–46, however, when the United States transferred from war to civilian industry $57 billion (27 percent of GNP) in one year, and without any serious unemployment, indicates that, given adequate aggregate demand, the adjustment capacity of the American economy is little short of fantastic.

QUESTIONS AND EXERCISES

1. "The main economic business of government is to act as a 'Governor'—i.e., to throw inflationary forces into the system when deflation is occurring and deflationary forces when inflation is occurring." Discuss and illustrate.
2. Distinguish between the "regulative" and the "positive" aspects of government intervention in economic life. Then discuss the following proposition: "It is the failure of government to fulfill its positive obligations which forces it to engage in all kinds of unnecessary regulative activities."
3. Define carefully as many meanings as you can which might apply to the word "inflation." Then discuss the following proposition: "Inflation is a deliberate swindle on the part of the government."

4. Under the Treaty of Versailles the German government had to pay a very large sum to the Allies in reparation payments. What do you think would be the effect of these payments on the foreign exchanges and on the course of international trade?
5. Suppose the government decided to extinguish the national debt by raising a capital levy, i.e., by taking industrial securities from their present owners in a "levy" and giving them to the present holders of the national debt in payment for government bonds. What effect would this measure have on the distribution of national income? Discuss the equity of such a proposal.
6. Discuss the following proposition: "The velocity of circulation is the one uncontrollable element in the monetary system. If we are to control money, we must find antidotes for changes in its velocity of circulation."
7. Suppose we have a system like that of Fig. 29, page 237, in which the private absorption function is $B = 35 + \frac{4}{5} Y_d$. The unit of income we suppose is \$1 billion.

   a. What will be the equilibrium level of private disposable income for deficits of 0, 5, 10, 15, 20, 25?

   b. Suppose now government absorption is 50. What will be (i) private disposable income, and (ii) total national income for total tax collections of 25, 30, 35, 40, 45, 50?

   c. Suppose now that government absorption rises to 100, and that the full employment level of total income at the current price level is 250. How much must be collected in taxes if inflation and unemployment are to be avoided?

   d. Suppose now government expenditure rises to 125, but that the practicable limits of the tax collection is 100. By how much will private absorption at the equilibrium level of private disposable income have to fall if total income is to be kept at 250?
8. In 1963 the United States Government carried out a tax cut as a remedy for unemployment. Discuss the theory behind this measure, and indicate what conditions have to be fulfilled for it to be successful. Illustrate by means of a mathematical model.

CHAPTER 11

POSITIVE AND NORMATIVE ECONOMICS

We have now reached the end of rather a long road. It may be appropriate, then, to take a brief look at the way we have come, and to ask ourselves whether these high abstractions have any real significance for man and his problems. Some economists, indeed, might deny themselves the pleasure—or pain—of asking the last question. We do economics, the Cambridge economist, Pigou, is supposed to have said, because it is fun. For those who can play and enjoy the game, it *is* fun—life has a few purer pleasures than the detection of a fallacy. For many of us, however, economics is more than fun; it is an intellectual pursuit of importance for the welfare of mankind. An enormous amount of poverty and misery might have been avoided in the past if the intellectual tasks of economics had been better performed and more widely understood. There is something to know here that hurts us when we do not know it, and because of this, economics cannot merely be motivated by idle curiosity.

Economists Have Always Prescribed

Nevertheless the distinction which many writers have made between positive and normative economics is valuable. Positive economics is the study of "what is"—the facts and the principles which govern the actual course of economic systems, without any attempt to judge whether one state of the system is any better or worse than another. Normative economics goes on to make judgments on whether one state of the system is better than another, and goes on from there to make prescriptions and gives advice in regard to policies to be followed. It is hard to find an economist who is so "pure" that he has made no normative propositions. Certainly the classical economists had no hesitation about making a strong case for free trade, or even against it! The Keynesian economists

have had no qualms about prescribing remedies for depressions or inflations. The development economists, too, show little hesitation about offering good advice to poor countries who want to get rich quick. Nevertheless, it remains a good idea to separate positive from normative propositions, and to be clear about the ethical assumptions which must underlie all normative propositions regarding society.

How Our Images of Fact and Value Are Interrelated

The relation between normative and positive propositions cannot simply be dismissed, however, by segregating them. In so far as normative propositions are accepted and influence human behavior, especially the behavior of powerful decision-makers, they will themselves have an effect on the nature of the positive system. That is to say, an essential part of the positive system is what people, in fact, believe about normative propositions. Similarly, normative propositions themselves, and what people believe about the goodness or badness of various states of the positive system, depend on our image of the positive system and its possible positions. We dismiss the unattainable grapes as sour, and our attitude toward such phenomena as unemployment, poverty, and war depends very much on whether our image of the possibilities of the positive system includes the possibility of eliminating these conditions. There is, therefore, a complex mutual interaction between our images of the positive and of the normative, and they are not always easy to separate. Their separation, however, is one of the intellectual tasks of analysis, and one that should not be shirked.

By the nature of its subject matter, economic analysis is peculiarly well fitted to apply its positive techniques to the analysis of certain normative problems, because of the fact that positive economics has to deal with the phenomenon of choice among goods as an essential part of its subject matter, and for normative economics also the problem of choice is fundamental. Positive economics may deal only with the choice between ham and eggs, and normative economics, with the choice between good and evil, and though these are not the same problem they are more similar than many moral philosophers like to think.

WELFARE THEORY

The Welfare Function

A normative statement takes the form, A is better than B. It implies, therefore, that the various alternatives in which we are interested can be *ordered* on a line so that any alternative on one side of any point

is regarded as better than any alternative on the other side of it. This is what is generally meant by a *welfare function*. In economics we can postulate a welfare function such that all alternative positions of the economic universe (i.e., all combinations of prices, outputs, incomes, and all other economic quantities) can be ordered or arranged in this manner. The combination or position at the head of such a list is the *optimum* position. All value judgments imply an ordering of this nature. Immediately, of course, we run into the difficulty that the welfare functions of different individuals are different. One individual may even have different welfare functions, depending on the role he is playing. He may say, "A is worse than B for me, but better than B for the country." Economics cannot say which of all possible welfare functions is itself the "best." It has been shown by Kenneth Arrow[1] that it is not even generally possible to add up or otherwise combine the welfare functions of different individuals into a "social welfare function" to which they will all voluntarily consent. Other social sciences may be able to say something about the way in which compromise social welfare functions are derived from discussion and political activity. The contribution of economics, however, lies mainly in the analysis of the formal properties and plausible general characteristics of welfare functions, especially those which involve economic variables. These properties and characteristics are similar to those of utility or production functions.

Value Analysis

We can, therefore, perform something with the tools of economic analysis which might be called "value analysis," and which should be of use in clarifying the choices involved in economic policy. It is necessary first to identify certain broad characteristics of systems which can be regarded as significant from the welfare point of view. These characteristics, or "subordinate ends," must be capable of rough measurement, and must be related to a welfare function in the sense that we can arrange the various values of the subordinate end in an order of "goodness." Thus we might distinguish the following subordinate ends of economic policy: (1) the rate of economic progress; (2) the short-run stability of employment, incomes, or prices; (3) equality in the distribution of income; (4) absence of waste in the allocation of resources among different industries and occupations; (5) width of freedom of personal choice in regard to commodities, occupations, or modes of life; (6) the degree of satisfaction with the human relations involved in economic processes. This list is by no means exhaustive; it covers, however, a good deal of the territory.

[1] Kenneth Arrow, *Social Choice and Individual Values*, 2nd ed., New York, Wiley, 1963.

Possibility Functions Relating Ends

The task of value analysis, then, is twofold. It is to examine the "possibility functions" which limit the achievement of these subordinate ends. We want to know, that is to say, what combinations of these subordinate ends are possible or compatible. We may not be able to plot the whole range of this function, but we need to know something about its marginal properties. We need to know, for instance, at least roughly, what we have to give up of one end in order to gain more of another. If, for instance, we achieve greater equality of income, does this mean that we must accept a slower rate of economic progress? If we increase the rate of economic progress, does this mean that we have to pay for this in a certain rate of inflation or in the restriction of individual choice or in the development of strained human relations and cultural disintegration? If we have wide freedom of personal choice, does this involve us in wider fluctuations of employment? The answers to these questions depend on the empirical nature of the system as it is; they do not involve any judgments about the welfare function, or what ought to be. Without some knowledge of the possibility function, however, our welfare judgments may be irrelevant. There is no point in wanting the impossible or in crying for the moon, and the essence of the intelligent value judgment is the selection of the best out of all possible situations.

Welfare Orderings of Ends

In order to find the best situation, however, we must have a welfare function, and even though we cannot specify that any particular welfare function is the "right" one we can explore certain general properties of welfare functions as a safeguard against partial and uninformed judgments. Thus we can assert with some confidence that no single subordinate end can serve as a measure of welfare as long as there are others which may compete with it or complement it. A "law of diminishing returns" is likely to apply to the welfare ordering of any *single* end, such as economic progress, if the quantity of other ends is held constant. If we suppose for a moment that "goodness" or social welfare, W, can be measured, then if $A, B \dots$ represent the different subordinate ends, we can postulate a social welfare function, in the mind of any one person.

$$W = F(A, B \dots) \tag{1}$$

With two subordinate ends, A and B, this could be represented by a set of indifference curves,[2] and a maximum social welfare found within the possibility boundary.

[2] See Volume I, Chapter 27.

A Social Welfare Model

A not uncommon pattern is shown in Fig. 31. We measure the subordinate ends, A and B along the axes, OA and OB. We suppose a possibility boundary, $A_0A_1B_1B_0$; combinations within this boundary are possible, outside it are not. Between A_1 and B_1 the two subordinate ends are competitive; we have to sacrifice one in order to get more

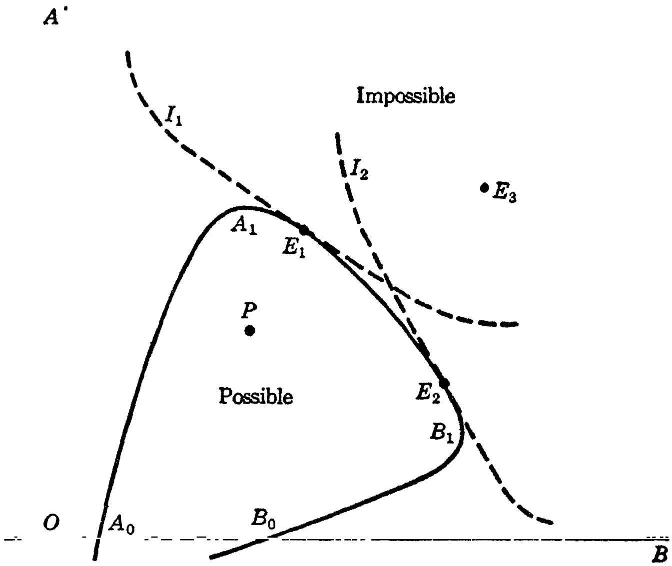


Fig. 31. Social Welfare

of the other. At some point, however, as we continue, say, diminishing B in order to get more of A, we find we can get less and less of A in return, and eventually, as at A_1, the curve reverses itself and diminishing B diminishes A also. That is, between A_1 and A_0, and between B_1 and B_0, the two subordinate ends are complementary. The rationale behind this assumption is that beyond a certain point the problem of general social disorganization becomes much more important than the specific ends—if failure to achieve a specific end becomes serious enough, the social disorganization which will result limits the achievement of all the other social objectives as well. Thus, suppose A measured the rate of economic growth and B, the equality in the distribution of income. Within the range, A_1B_1, we suppose these may be competitive, in the sense that once we have reached the possibility boundary, a higher rate of growth can only be achieved by making incomes less equal. If we

diminish equality beyond the point A_1, however, we lose rather than gain in economic growth, moving back along the line from A_1 towards A_0. Similarly, as we lessen the rate of economic growth in the hope of achieving greater equality, we may find beyond B_1 that this hope is frustrated and at low levels of economic growth equality becomes more difficult to achieve.

The Sensitivity of the Preferred Position

If, now, we can postulate a set of social welfare indifference curves, one of which, such as I_1E_1, touches the possibility boundary at E_1, then E_1 is the preferred position within the constraints of the possibility function. The slope of the indifference curves is a measure of the relative value of the two subordinate ends in the mind of the evaluator. Thus, if in Fig. 31 the indifference curve is flat, this means that the evaluator would be willing to sacrifice a lot of B to get one unit of A—that is, he values A very highly and B, little. Similarly, if the indifference curves are steep, it means that the evaluator cares little for A and a lot for B. If the evaluator cares only for A and is quite indifferent to B, his indifference curves will be horizontal and his preferred position will be A_1. If he cares only for B his indifference curves will be vertical and his preferred position will be B_1. One important conclusion is that if the possibility boundary itself has a fairly constant slope in a certain range, a small change in the preference or welfare function may cause a large change in the preferred position. Thus a relatively slight rise in the preference for B, as reflected in the shift from indifference curves of type I_1E_1 to curves of type I_2E_2, causes a sharp shift in the preferred point from E_1 with a lot of A and a little of B, to E_2 with a lot of B and a little of A. We cannot judge the strength of preferences, then, merely by the position of the preferred point, unless we know something about the image of the possibility function. If E_1E_2 were a straight line, then a very small shift in preferences could cause a shift from E_1 to E_2 or from E_2 to E_1. Thus it is possible that the difference between, say, socialist and capitalist countries represents quite a small difference in basic values, which results, however, in a large difference in their preferred position. Thus suppose in Fig. 31, B measured the degree of socialism, say the proportion of the economy in the government sector, and A measured the amount of personal liberty, measured perhaps by the proportion of the economy in the internal police and security forces. There may well be a fairly linear possibility boundary here between, say, E_1 and E_2, with increasing amounts of socialism involving relatively constant decreases in personal liberty. Under these circumstances a relatively slight preference for socialism may move the society to E_2 as its preferred point, and a relatively slight preference for personal liberty may move it to E_1.

Social Development

The question as to whether a society can be at a position inside the possibility boundary, say at P, is interesting, but is perhaps mainly a semantic question. We may be within the possibility boundary because of ignorance; in this case, however, ignorance itself is the boundary, and it can be argued, therefore, that if we are not on the possibility boundary, we will always move to it. The real problem here is whether resources can be devoted to the task of "pushing out" the possibility boundary itself so that we can have more of *all* the subordinate ends. This is the meaning of "development"; just as economic development enables us to have more of all commodities, so social development means that we can have more of all of our various subordinate ends. Just as inventions in the technology of production enable us to push out the production opportunity line, so "social inventions" enable us to push out the possibility boundary of ethical goals. The invention of religious toleration, and the separation of church and state, for instance, has enabled us to combine many religious goals in a single society, whereas previously it was thought that all members of the same society must practice the same religion.

Conflict and Development

The ability of a society to devote resources to development, whether economic or social, depends to a considerable degree on its ability to handle conflict. The existence of a possibility line, like E_1E_2, implies conflict among goals in the sense that if we achieve more of one, we will achieve less of the other. If these various goals are embodied in parties, factions, or classes within the population, the conflict among goals may result in internal conflict within a society, or external conflict between societies which hold diverse subordinate goals. Conflict, however, is costly in resources, and if many resources are devoted to conflict, not many can be spared for development, that is for pushing out the possibility boundary. Thus we have a curious dilemma; at low levels of development, where scarcity either of commodities or of social goals is keenly felt, conflict is likely to be intense and hard to manage, and may indeed absorb so many resources that none can be spared for development, and the society stagnates. Once development starts, however, conflict tends to become less intense, for when there is more for everybody there is less incentive to gain at the expense of someone else. As conflict becomes less intense, however, more resources can be spared for development, development proceeds faster, conflict becomes still less intense, and so we emerge into an almost irreversible and cumulative process of growth. It may well be, therefore, that social inventions which diminish the resources devoted to conflict may well be the initial impetus which

starts the cumulative process of economic development. In Britain, for instance, it is significant that the political settlement of 1688, which brought to an end half a century of intense conflict, seems to have opened up the way for an era of internal peace and great economic development.

GROWTH AS A MAJOR OBJECTIVE

Growth and Equality

In assessing the various subordinate ends of economic policy, it is difficult to avoid giving first place and a very high priority to growth and development, not only because of their intrinsic importance, but because they tend to be highly cooperative in the long run with the achievement of other goals. Thus equality can only be afforded by rich societies—in a poor society equality would condemn everyone to a miserable subsistence and the society to stagnation. A society of uniform mass poverty can hardly hope to accumulate or to concentrate resources in the hands of the innovators. At first, economic growth frequently seems to accentuate inequality, as those few who participate in it draw away toward higher incomes from the mass of the society which still remains in the primitive condition. The end result of economic growth, however, seems to be a more equal distribution at the higher level, when all members of the society have been caught up in the more productive culture. As the pioneering middle class moves out of the morass of universal and equal poverty, inequality increases; but if the progress continues, the whole society is gradually drawn up into the plateau of universal and relatively equal comfort—a condition which we would be happily approaching in the more advanced countries, were it not for the specter of war.

The Attack on Economic Rent

A premature attempt at too great an equalization of incomes, then, may destroy productive activity and may hamper economic growth by taking away its rewards. Out of our analysis, therefore, comes a piece of advice to the lawmaker: wherever possible, attack economic rent. Economic rent is defined as any payment to the owner of a factor of production in excess of what is required to keep that factor in continuous service. If such an excess is absorbed by taxation or regulation, there will be no diminution in the quantity of the factor supplied. But if taxation digs deeper than this surplus and attacks the actual "supply price" of the factor, the quantity supplied will be adversely affected. The advice is good; the difficulty is, of course, to find the economic rents. We certainly

cannot assume either that economic rent is peculiar to the income from land or that the income from land is all economic rent. A great deal of labor income is economic rent, and in part the services of land are a "produced" commodity with a quite definite cost. We cannot even assume, with Henry George, that increments in land values are "economic rents." The increment of land value may be in many cases the reward of pioneering; and if full taxation of land values had been in effect in the 19th century, it is probable that the great westward expansion of the United States would never have taken place, for the expectation of rising land values was one motive which inspired the pioneers to invest their lives and comforts in the development of an empty continent. It is probable that a properly constructed income tax falls to a very large extent on economic rents. In so far as it applies to all occupations it does not affect *relative* profitabilities, and so cannot be escaped by shifting occupations. It may have some effect in reducing activity where the amount of labor offered by an individual is within his control, especially where the marginal rates of taxation are high. It may have some effect in discouraging risky innovation, especially where this may lead to highly irregular incomes, as the high tax rates of the successful years are not offset by the losses of poor years. There does not seem to be much direct evidence, however, that these effects are serious at the degree of progressiveness which prevails in the American tax system, though in so far as the deleterious effects are cumulative, they may not show up in the short run.

Growth and Misallocation

Economic progress, as we have seen, inevitably involves the redistribution of resources among various occupations. Such redistribution, however, is inevitably accompanied by losses to the owners of resources, whether capital or labor, in a declining occupation and gains to those in an expanding occupation. The best solution to this problem is mobility—that is, rapid redistributions of resources in accordance with new structures of demand. Sometimes, however, mobility may not be possible, or may be costly, as in the case of highly specific equipment or skills. Such unavoidable losses may properly be charged as part of the cost of economic progress, and it seems unfair to make those on whom they normally fall bear the whole burden. If, however, these costs are charged to the innovator, we run grave risks of suppressing progress altogether. Consider, for instance, the philosophy of "equalizing competitive advantage"—a theory which has been applied, for instance, both to the tariff and to the taxation of chain stores. Tariffs are sometimes justified on the

grounds that the tariff should compensate for the lower costs of the foreign producer, and chain store taxation is justified on the grounds that chain stores have lower costs than independents. Such a doctrine applied at an earlier date would have taxed railroads to the point where stagecoaches could compete with them, power looms to the point where hand looms could compete with them, and even spades to the point where fingers could compete with them. This is not to say that society has no duty to those dispossessed by technical progress or by changes in demand. But its efforts should be directed toward increasing mobility rather than toward freezing the existing structure by taxes and subsidies, and if compensation is to be given, it should be charged to society as a whole.

The problem of evaluating the market structure of a society—that is, its competitive, monopolistic, or oligopolistic nature—is one of great difficulty. From the point of view of allocation and distribution there is a strong case against monopoly.[3] Too much competition, on the other hand, may be inimical to economic progress, for unless there is some degree of monopoly protection the fruits of innovation cannot be enjoyed by the innovator. This is the justification usually given for patent and copyright laws. Similarly, there are wastes of monopolistic competition in terms of undercapacity operation and competitive advertising, but there are also gains in terms of opportunities for variety and decentralization of control of information.

Growth and Economic Freedom

Economic freedom, like equality, is a subordinate end which seems, over a considerable range, to be complementary to economic growth. It is a vague concept capable of many meanings. In one sense freedom is virtually synonymous with "power," and in this sense economic freedom is largely synonymous with "riches." The poor have very little freedom to do anything, and if freedom is the ability to do things and go places, the richer we are the more freedom we have. On the other hand, there are other meanings and dimensions of freedom. Freedom in the sense of power is measured by the area of our possibility field—the more things possible to us, the greater our freedom. Freedom in the sense of the rule of law—that is to say, freedom from arbitrary coercion or limitations imposed by threat or the arbitrary will of another—is a matter of the nature or "quality" of the boundary: what it is that fences us in, rather than how big is the corral! Even in this sense freedom is easier to achieve in rich societies, for in a poor society coercive rule is almost a necessity if any economic surplus is to be squeezed out of the masses. A good case can be made, therefore, for the assumption that freedom, like equality, is

[3] See Volume I, Chapter 23.

a "rich man's good" and only comes widely with economic growth that makes us *all* rich!

Growth and Stability

It is still a somewhat unresolved question whether economic stability is competitive or complementary with growth; this is again perhaps a place where social invention is of great importance. In the economic development of the West, growth seems to have involved a good deal of instability—the long-term growth has not been steady, but has come in fits and starts, with rapid growth during the boom and little growth, or even retrogression, during depression. What we want to achieve by a stabilization policy, of course, is steady growth; we certainly do not want stability at the price of stagnation. It is still something of an open question whether an occasional "shakedown" or interruption in growth, by getting rid of inefficient firms and getting resources into the hands of new entrepreneurs, may not result in a faster overall rate than a steady growth rate could provide. It is certainly significant that economic growth in the socialist countries has also been interrupted by periods of hesitation or even retrogression, like the first collectivization in the Soviet Union (1928–32) and the sadly misnamed "great leap forward" in China, which led to some years of stagnation or even retrogression after 1959. Nevertheless, we should not assume too easily that rapid steady growth is impossible—certainly the case of Japan, which has probably grown more steadily than any other economy, both during the period 1880–1930 and the period after 1945, indicates that rapid growth is consistent with much greater stability than most Western countries have achieved. Here again we must not underestimate the role of social invention, and the task of developing social institutions which permit rapid and steady growth is not intrinsically impossible.

Economic Growth and Human Character

Finally we come to the most difficult problem in evaluation, and yet one which cannot be shirked in the total evaluative process. What is the effect of various forms of economic life and institutions on individual character and personality? Perhaps all that the economist dare do here is to point to the problem. Does the division of labor, as even Adam Smith suggested, produce narrow, ill-rounded characters? Does the increase of power which economic progress brings merely enable us to damn ourselves all the more easily—to satisfy the evil lusts and ambitions from which poverty would keep us free? Does the increase in the scale of organization which results from economic development impair personal freedom and the intimacy of personal relationships? These are questions

to which the economist as such can give no clear answers, but which are nevertheless involved in any final appraisal of economic institutions and policies.

Conclusion: The Dismal Science?

The conclusions of our analysis will be disappointing to the student who seeks a panacea for all our problems. In the days of Malthus, economics earned justly the title of the "dismal science," not because it was dull, but because its conclusions were so depressing to those who looked for a bright future for humanity. The Malthusian specter still haunts us, perhaps even in the developed economies. But the Malthusian hope also remains—that if it is expensive tastes and a rational ethic rather than the biological urge which limits the growth of population, a stable and rich society is possible. We are still in the thick of the technical revolution, and where it is carrying us no man can say; but if it does not take us to destruction, there is at least a reasonable hope that it may take us to a new level of society as far removed from past and even present civilizations as they were from the barbarism which preceded them, a society in which war, poverty, and disease will be effectively eliminated.

There is no cause, however, for undue optimism. If the dark dynamics of population growth does not dash the cup of plenty from our lips, the darker dynamics of national and ideological warfare may do even worse things to us. The hope which many once saw in socialism and in the conscious control of man's destiny has been dimmed by the dreadful spectacle of the Communist world—the liquidation of the kulaks, the reliance on slave labor, the shocking disintegration of all decency and simplicity in personal relationships, the monstrous corruption and tyranny of the one-firm state. There is a dismal science of politics as well as of economics, and the attempt to abolish the free market and to set up a planned economy runs into economic and human diseconomies of scale, the substitution of violence and coercion for economic advantage as the prime mover of men, and the unspeakable horror of the manipulative society.

On the other hand, it must not be thought that the economist brings uniformly cheerful tidings to the supporters of capitalism. His results are profoundly disturbing to any who believe that capitalism with its attendant democracy is preferable to any form of planned economy. The optimism of *laissez faire,* if it ever existed, is gone. It has become clear that an unregulated capitalism is liable to serious disorganization; it is subject to business cycles and periodic unemployment. It may also suffer from more serious ills: from "secular stagnation," as opportunities for further accumulation get used up, or from a creep-

ing paralysis of political control and monopolistic restriction. It suffers perhaps even more from certain sociological diseases: from a failure to satisy the deep needs in man for community, for "belonging to" something greater than himself. Hence, nationalism in extreme forms steps into the emotional gap which capitalism creates, and by the very growth of state power the foundations of a society based on a multiplicity of private institutions is destroyed.

We must recognize both the contribution and the limitations of economic analysis. Without a grasp of what it has to offer, we shall miss our way among the intolerable complexities of social life and the labyrinthine paths of history. But of itself it does not claim to provide a complete interpretation of history; we cannot say from economic analysis alone whether, for instance, socialism is inevitable or whether capitalism will survive. It does, however, point toward a certain philosophy of history and toward certain possible lines of human development. It encourages the hope that we may be able to find, somewhere between the ruthless extremes of totalitarian, monolithic socialism, and rudderless capitalism a developing middle way toward a "governed economy," in which the state shall have certain clear and perhaps extensive economic functions to perform in its capacity as a "governor" of the system, but in which the virtues of a "polylithic" society of many diverse and interacting free institutions and free individuals may be preserved.

Towards a Modest Optimism

The optimistic economist, therefore, is entitled to a certain faith that his science does indeed have immense consequences for the future welfare of mankind, and that even its past achievements, imperfect as they are, have made their mark. When we contemplate, for instance, the striking difference between the economic history of the 20 years from 1919–39 with the Great Depression, low rates of growth, and stagnant colonial empires, and the 20 years from 1945–65, with no serious depressions, much more rapid growth rates, and the dissolution of the old empires, it is hard not to give some credit for this difference to the progress of economics. We still, of course, face the cold war, and the positive probability of fatal nuclear disaster, which hangs over our heads. Even here, however, economics can modify, and is modifying, the intensity of the conflict. As Marxism takes its place as a special case of a more general social theory, as the socialist countries come to understand the value of free markets and the importance of relative prices, and as capitalist countries come to realize the need for a wise governance of the overall economy and a clear image of the future, we may not

unreasonably hope to see a diminution of the present ideological struggle until ideology is finally swallowed up by advancing knowledge. In this process economics will play a highly significant part, until it, in turn, perhaps is absorbed in a greater and more unified social science.

QUESTIONS AND EXERCISES

1. Discuss the possible shapes and properties of the welfare and possibility functions, as in Fig. 31, for the following variables:
   (a) A = the probability of peace, B = the probability of victory
   (b) A = Justice, B = Freedom
   (c) A = Security, B = Progress
   (d) Any other variables that interest you!
2. Extend the analysis of Fig. 31 to the case of three subordinate ends in three dimensions, *A, B,* and *C.*

APPENDIX

THE LITERATURE OF ECONOMICS

The body of economic analysis presented in this work is the result of a long and often painful process of thought on the part of many minds through the years. The student of economics cannot appreciate adequately the significance of present-day theory unless he has a substantial acquaintance with the great works of the past and with the history of ideas. The study of past errors is a useful discipline against present errors. The study of how the peculiar circumstances of a past time led its thinkers to ascribe general validity to particular institutions serves as a constant warning against ascribing too much generality to the peculiar circumstances of our own day. Any one author is limited by the scope of his own interests, and the student can never be acquainted with the scope of his subject unless he approaches it through the mediation of many different minds. It is not the purpose of this appendix to outline a systematic history of economic thought, nor an exhaustive bibliography of the literature; rather, is it intended to give some indication to the student of the points of origin of the principal ideas he has encountered in this work, and to guide his interests in further reading. There are several excellent histories of economic thought;[1] and the student would be well advised to read at least one of them, to get the broad historical picture.

Books about great books, however, are no substitute for the great

[1] Alexander Gray, *The Development of Economic Doctrine,* Longmans, Green, 1947; Erich Roll, *A History of Economic Thought,* Prentice-Hall, 1942; Edmund Whittaker, *A History of Economic Ideas,* Longmans, Green, 1940; R. L. Heilbroner, *The Worldly Philosophers,* Simon and Schuster, 1953; T. W. Hutchinson, *A Review of Economic Doctrines,* Oxford University Press, 1953; Joseph A. Schumpeter, *History of Economic Analysis,* Oxford University Press, 1954; Henry W. Spiegel (ed.), *The Development of Economic Thought,* Wiley, 1952.

books themselves, and the serious student should not long delay the study of the classical works of theory. The greatest of these, even after more than a century and a half, is Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations* (1776).[2] Lacking virtually all the mathematical and graphic techniques of modern value theory, Adam Smith nevertheless perceived, not always with perfect clarity but always with astonishing insight, the essential relationships that lie at the heart of economic life. What I have called the principle of equal advantage, and the idea of the movement of resources under the stimulus of prices and profits from one occupation to another, which is the central idea of the theory of value and distribution, stems directly from Adam Smith. Chapters 7 and 9 of his Book 1 could be used today with hardly any modification in an accurate elementary text. In his concept of "effective demand" (in modern terminology, the quantity demanded at the normal price) lies the germ of the modern theory of supply and demand. His distinction (unfortunately named, but vital nonetheless) between "productive" labor that was embodied in goods and "unproductive" labor not so embodied was the foundation of the theory of capital. His great defense of free trade against the mercantilists in Book 4 is famous, but it must not be thought that he was a bigoted advocate of "laissez faire." Indeed, almost half the work (Book 5) is a discussion of the proper functions of the state. Then his Book 3 outlines with great insight a theory of economic progress, previously much neglected, but now, with the widespread interest in economic development, worth careful study. It is indeed almost depressing to go back to Adam Smith on economic development and see what little progress we have made on the subject since 1776! Most of all, the *Wealth of Nations* is worth reading for its style and spirit—full of wisdom and the observation of life, universal in the material from which it draws. Of Adam Smith it may often be said that he draws correct conclusions from faulty reasoning, his insight and wide observation affording him conclusions which his imperfect analytical techniques did not enable him to prove. More than any other economist, with the possible exception of P. H. Wicksteed, Adam Smith embodies the liberal tradition of humane letters. The student who learns to love his quiet wit and keen but gentle observation of humanity will never degenerate into a narrow-minded specialist.

A writer of very different character, yet in his own way almost equally important, is David Ricardo. The student should read his

[2] Cannan's edition (Methuen, 1904) is probably the best. This has been republished in the Modern Library.

Principles of Political Economy and Taxation (1817)[3] not for its style, which is arid and humorless, nor for its conclusions, many of which are erroneous, but for its crystalline logic. Ricardo's mind cut through the inconsistencies of Adam Smith and reduced economics to a system in which the conclusions followed inexorably from the axioms. Unfortunately, the axioms were not always correct, and the conclusions suffer accordingly, but the student who follows the argument will have an excellent training in the discipline of economic logic. It is to Ricardo that we owe the first faint beginnings of the marginal analysis, in his statement of the law of diminishing returns and his theory of rent. The modern theory of the economic surplus arising in the case of a less than perfectly elastic supply had its beginnings in the Ricardian theory of rent, though its modern applications are much wider than Ricardo imagined.

In reading the classical economists it must constantly be borne in mind that their system is a special case of our more general modern constructions. The relatively little importance given to demand in their theory of value may be traced to their general (implicit) assumption that supplies are perfectly elastic. In such a case, of course, prices do not in the long run depend on demand at all, but are determined by the level at which the supply is perfectly elastic, which in its turn is determined by "the cost of production." Ricardo was perhaps the first to recognize the possibility of what we now call an imperfectly elastic supply in the case of land, and saw that a rise in demand would raise the price at least of foodstuffs by pushing production on to poorer lands. Even Ricardo, however, assumed implicitly a perfectly elastic supply in the case of labor in his subsistence theory.

Ricardo and Torrens between them developed the theory of comparative advantage in international trade, though it was left to an economist of the next generation (Cairnes) to perceive the generality of this principle. With Ricardo, indeed, "model building" enters economics fairly explicitly, even though it is implicit in Adam Smith, and Ricardo does not go beyond the mathematical equipment of arithmetical examples.

Another writer of Ricardo's time who is worth reading today is the Rev. T. R. Malthus. He is principally famous for his essays on population, the first of which, 1798,[4] is a brilliant essay in moral theology on the theme of the long-run equilibrium of human misery; but it does not carry economics much beyond where Adam Smith left it. In the

[3] The magnificent edition of Ricardo's complete works in nine volumes by Piero Sraffa (Cambridge University Press, 1955) is the definitive version. The *Principles* are available in Everyman's Library.

[4] The *Essay on Population* is in Everyman's Library and also in Ann Arbor Books.

light of modern monetary theory, however, his *Principles of Political Economy* (especially Part II) stands out as a work of remarkable insight, which anticipates many of the Keynesian principles and yet was largely neglected for over a century.

Another writer whose true significance did not appear till a later date was A. A. Cournot (1801–1877), whose *Recherches*[5] (1838) developed, in mathematical form, much of the modern theory of the firm; the concepts, though not the names, of the marginal analysis were first developed by him. He should be read by any student who has some elementary knowledge of the calculus. The mathematical form of his writing, however, prevented him from having much influence on the economists of his own time, and the classical system, especially as expounded by J. S. Mill,[6] held undisputed sway until about 1870.

Kark Marx (1818–1883) stands in a class by himself. *Das Kapital*[7] (1867) is a book which has had a profound influence on the world. It represents an early attempt to develop a theory of the economic system as a whole and of its progress in time. In this attempt, however, Marx was greatly handicapped by the inadequate analytical apparatus which he inherited from Smith and Ricardo; consequently, errors which can be generously interpreted as matters of exposition in the classical economists, such as the labor theory of value, are erected into the foundation of a logical but inadequate system of economic reasoning. Although Marx made little contribution to the broad line of development of economic thought, the student should read at least the first volume of *Das Kapital* or, better still, Borchardt's condensation. To study the errors of a great, if wrongheaded, mind is often more valuable than to skim the platitudes of a small one.

The next important group of writers constitute the so-called "marginal utility school." The ideas of this school were developed at about the same time (c. 1870), independently in England by Stanley Jevons, in Austria by Karl Menger and Friedrich von Wieser, and in France by Léon Walras. With this school began the extended use of mathematics in economic analysis, foreshadowed by Cournot; for though many of its exponents expressed their ideas in literary form, their theories were essentially mathematical in structure. Perhaps the great-

[5] *Recherches sur les principes mathématiques de la théorie des richesses,* English trans. by N. T. Bacon, Macmillan, 1897; this contains also a useful bibliography of early mathematical economics.

[6] *Principles of Political Economy* (1848).

[7] The translation by Eden and Cedar Paul (Allen and Unwin, London, 1928) is very good; the Modern Library edition, containing Stephen Trask's translation of Borchardt's condensation, together with "The Communist Manifesto," is an excellent introduction.

est work of this whole school is P. H. Wicksteed's *The Common Sense of Political Economy*,[8] even though it appeared forty years after the first formulation of 1870. This is a book which every student of economics should read, in spite of a certain prolixity and an occasionally labored style. It is couched (often at the cost of being cumbersome) in nonmathematical language, and it provides the most consistent and highly developed exposition of the utility analysis as the foundation of both demand and supply, and as a general theory of choice. Perhaps Wicksteed's greatest contribution was his demonstration that economics is not merely a matter of the market place or of financial dealings, but is one aspect of *all* human activity—namely, the aspect of choice, or the balancing of alternatives one against another, where limited means have to be apportioned among competing ends. His delightful discussions of how much family prayers should be shortened to speed a parting guest to the train, or of the value of a mother-in-law in terms of how high a cliff one would drive off to save her, should open the eyes of every student to the great *generality* of economic principles. Wicksteed, more than any other, laid the ghost of that shadowy creation, the "economic man." Wicksteed also made valuable contributions to the generality of economic principles themselves, and showed, for instance, the essentially similar derivation of supply and demand, and the universal character of the "law of diminishing returns." The concept of a homogeneous production function, and the distinction between diminishing returns to various proportions of factors and diminishing (or increasing) returns to scale, also owe much to Wicksteed.

The student whose acquaintance with mathematics extends to the elementary calculus should read Jevons's *The Theory of Political Economy*, in spite of the many errors which it contains. Even the nonmathematical student will find much of interest in this work; its freshness of style, its enthusiasm, the sense of discovery, of the opening up of vast new areas of conquest to the human mind, can hardly fail to be inspiring even if, as Marshall pointed out, Jevons was much less of a revolutionary than he himself thought.[9] Much of what the marginal utility school proclaimed is implicit in the classical economists, and, indeed, almost became explicit in the work of Nassau Senior, a contemporary of Ricardo, and was made quite explicit in the unnoticed work of the tragic Gossen (1854). It was not, however, until the productive 1870's that the dependence of value on scarcity in relation to demand, and of scarcity on cost of production, was made quite clear. Von Wieser,[10] in particular, helped to clarify the rela-

[8] Fourth ed., Macmillan, London, 1910 (reprinted 1924).

[9] See Alfred Marshall, *Principles of Economics*, 8th ed., Macmillan, 1938, appendix I, pp. 813-821.

[10] F. von Wieser, *Social Economics*, trans. by A. Ford Hinrichs, Greenberg, 1927.

tionships between the value of finished goods and of factors of production, and showed how the value of factors of production depended on the value of the goods which they produced.

It remained for Walras,[11] however, to bring together the complex relationships of economic life into a single mathematical system of mutual determination and mutual interaction. Walras is the Laplace of economics; just as Laplace transformed astronomy from a system in which the movement of each heavenly body was attached to its own particular cause to a system in which the mutual interactions of all bodies upon each other determine the behavior of all, so Walras transformed economics from a system in which each value was attached to its own particular cause to a system in which all values, whether of finished goods, intermediate products, or factors of production, are mutually determined by the interaction of the innumerable forces of desire upon the innumerable resistances of scarcity. The student who for want of mathematical ability cannot read Walras suffers under a severe handicap, and no better investment for the future economist could be recommended than the improvement of his mathematical skills to this point.

The marginal utility school is also important for its contributions to the theory of capital, though the exact significance of these contributions is still somewhat a matter of dispute. The basic idea of their theory of capital is that of a "period of production" between inputs and outputs. Because of this period of production inputs of "original factors" are embodied in intermediate products, which are "real capital." The volume of these intermediate products clearly depends on the length of the period of production. Jevons developed these ideas in a somewhat crude form, but their greatest development is due to Eugen von Böhm-Bawerk, whose *Positive Theory of Capital*[12] is essential to the study of this part of the subject. Another important contributor to this part of the subject was Knut Wicksell (1851–1926), whose little book *Über Wert, Kapital und Rente* (Jena, 1893) systematized the ideas of Böhm-Bawerk and by reducing them to mathematical form brought out their underlying assumptions.

The next great writer in point of time is Alfred Marshall, who may not unjustly be regarded as the father of modern Anglo-Saxon economics. Perhaps his greatest contribution is the development of supply and demand curves and of the concept of elasticity of demand and supply. These tools of analysis are so essential to the modern economist that

[11] Léon Walras, *Eléments d'économie politique pure*, 4th ed., Lausanne, 1900, trans. by William Jaffé, *Elements of Pure Economics*, Irwin, 1954.

[12] Trans. by William Smart, Stechert, 1923.

he is apt to forget their relative youth. In his *Principles of Economics*[13] Marshall developed a system not unlike that of Walras in essential principles but written in geometrical rather than algebraic language—losing thereby in generality but gaining in practicality. The basic proposition of the system is that the equilibrium output of a commodity is that at which the supply price and the demand price of the output are equal (Volume I). Marshall showed how the demand curve depended on the underlying utility relationships, and how the supply curve was related to costs. We are indebted to him for the distinction—useful, if dangerous —between short-run and long-run situations. We owe to him also certain developments of the theory of the firm in monopoly, though he did not succeed in integrating the theory of the firm explicitly into the main body of analysis.

Marshall began an era of great proliferation of economic writings in which we still live and in which, therefore, it is difficult to assess enduring values. As far as the tools of analysis are concerned, we have seen the development of the indifference curve, especially by Pareto, whose *Cours* and *Manuale*[14] are essential reading for the advanced student. In the theory of capital and interest the works of Irving Fisher are to be highly recommended.[15] H. J. Davenport is worth reading for his original point of view and his concept of alternative cost.[16] F. H. Knight's *Risk, Uncertainty, and Profit,* 1921, reprinted by the London School of Economics, 1946, is an important if somewhat unclassifiable work.

Perhaps the most important developments of twentieth-century economics lie in four fields. (1) There has been a great advance in what used to be called the theory of money, but now is frequently called "macroeconomics"—i.e., the theory of the broad averages and aggregates of the whole economic system, such as the general level of prices, of output, of wages, of employment, of interest rates, and so on. (2) There has also been an advance in extending the theory of relative values to include cases of imperfect competition and monopoly. We should perhaps include in this area, though the developments are somewhat unrelated, some considerable advances in the theory of the firm, both through the applications of the theory of linear programming and through the development of more dynamic, behavioral theories of the behavior of economic organizations. (3) Especially in the past ten or fifteen years there has been a great

[13] Eighth ed., Macmillan, 1938.

[14] *Cours d'économie politique,* Lausanne, 1896; *Manuale di economia politica,* Milan, 1906.

[15] *The Nature of Capital and Income,* Macmillan, 1906; *The Theory of Interest,* Macmillan, 1930.

[16] *The Economics of Enterprise,* Macmillan, 1913; *Value and Distribution,* University of Chicago Press, 1908.

revival of interest in the study of economic development, with some new theoretical contributions to dynamic economics. (4) The application of mathematical models, and more recently of computer techniques to economics, has made great strides, especially in the direction of handling models with large numbers of variables, for instance, input-output models.

Economists up to and including Marshall had many observations regarding the theory of money, but with the possible exception of Ricardo—whose cost-of-production theory of the value of money, however inadequate, at least fitted in with his general system—the theory of money was never integrated into the general theory of value. Although Adam Smith set out on an "inquiry into the nature and causes of the wealth of nations," the attention of economists in the 19th century became increasingly concentrated on the problem of "value"—i.e., of relative prices. The reawakening of interest in the broader problems of macroeconomics begins perhaps with Wicksell, whose *Geldzins und Guterpreise*[17] should be read by every student, and whose *Lectures on Political Economy*[18] are well worth the attention of the more advanced student. Irving Fisher's *The Purchasing Power of Money* is also a landmark, for it gives the first clear formulation of the "equation of exchange" ($MV = PT$). The modern developments in macroeconomics, however, are most closely associated with the work of J. M. Keynes (the late Lord Keynes). His early *Tract on Monetary Reform* (Harcourt, Brace, 1924) is still worth reading, in spite of some out-of-date material. It is an eloquent plea for the stabilization of prices as the object of monetary policy, and contains, along with D. H. Robertson's little book entitled *Money,* the essence of the "oral tradition" on this subject at Cambridge, England, which flowed from the teaching of Marshall. Keynes's major works are the *Treatise on Money* (Harcourt, Brace, 1931) and the *General Theory of Employment, Interest, and Money* (Harcourt, Brace, 1936). These are not easy to read, and are often confused in thought. Nevertheless, the serious student cannot fail to derive substantial benefit from their study; they open up vistas of intellectual exploration which have by no means been fully covered.

The theory of imperfect competition has been developed in two principal works—E. H. Chamberlin's *The Theory of Monopolistic Competition* (Harvard University Press, 1933) and Joan Robinson's *The Economics of Imperfect Competition* (Macmillan, 1934).[19] The elementary theory of the firm and industry in its modern form, especially as it involves the use of the marginal revenue curve, is to be attributed mainly to

[17] Trans. under the title of *Interest and Prices* by R. F. Kahn, Macmillan, 1930.

[18] Trans. by E. Classen, New York, Macmillan, 1934.

[19] Another important work in the field is Robert Triffin, *Monopolistic Competition and General Equilibrium Theory,* Harvard University Press, 1940.

these writers. The assumption of perfect competition which underlay so much of the classical economics, and even the Marshallian system, was shown by these writers to be a special case of a more general theory.

The theory of economic decision-making has received much attention in recent years. One aspect of this is the development of the theory of games. The student who finds the classic work of J. von Neumann and Oskar Morgenstern[20] heavy going may find a clear exposition in R. O. Luce and H. Raiffa.[21] Another aspect is "operation research," of which the theory of linear programming is a part; a brief résumé of these developments will be found in K. E. Boulding and W. A. Spivey.[22] Another development is in the analysis of dynamic patterns of decision-making and their computer simulation; R. M. Cyert and J. G. March[23] and the group at the Carnegie Institute of Technology have pioneered in this work.

Recent contributions to the theory of economic development are A. O. Hirschman[24] and H. Leibenstein;[25] this field seems almost to be moving out of economics into a general social science with such writers as Everett E. Hagen[26] and David C. McClelland.[27]

In econometrics the great pioneers were Henry Schultz[28] and Ragnar Frisch,[29] and Wesley Mitchell.[30] The name of Wassily Leontief[31] is particularly associated with input-output analysis. G. Tintner[32] and Stephen Valavanis-Vail[33] have written good textbooks.

Among more general works, those of J. A. Schumpeter[34] are important

[20] J. von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior,* Princeton University Press, 1944.

[21] R. O. Luce and H. Raiffa, *Games and Decisions,* New York, Wiley, 1957.

[22] K. E. Boulding and W. A. Spivey, eds., *Linear Programming and the Theory of the Firm,* New York, Macmillan, 1960.

[23] R. M. Cyert and J. G. March, *A Behavioral Theory of the Firm,* Englewood Cliffs, N.J., Prentice-Hall, 1963.

[24] A. O. Hirschman, *The Strategy of Economic Development,* Yale University Press, 1958.

[25] H. Leibenstein, *Economic Backwardness and Economic Growth,* New York, Wiley, 1953.

[26] E. E. Hagen, *On the Theory of Social Change,* Homewood, Ill., Dorsey Press, 1962.

[27] D. C. McClelland, *The Achieving Society,* Princeton, N.J., Van Nostrand, 1961.

[28] Henry Schultz, *The Theory and Measurement of Demand,* 2nd ed., University of Chicago Press, 1957.

[29] Ragnar Frisch was the first editor of *Econometrica* (Vol. 1, 1933).

[30] His last work, *What Happens during Business Cycles,* New York, National Bureau of Economic Research, 1951, is probably the best introduction.

[31] W. W. Leontief, *The Structure of American Economy, 1919–1939,* 2nd ed., New York, Oxford University Press, 1951.

[32] G. Tintner, *Econometrics,* New York, Wiley, 1952.

[33] S. Valavanis-Vail, *Econometrics,* New York, McGraw-Hill, 1959.

[34] *The Theory of Economic Development,* trans. by Redvers Ouie, Harvard University Press, 1935; *Business Cycles,* New York, McGraw-Hill, 1939; and *Capitalism, Socialism and Democracy,* New York, Harper & Row, 1942

for their insight into business cycles and into the whole process of economic development. The student of the dynamic process will find important contributions in J. R. Hicks's *Value and Capital* (2nd ed., Oxford University Press, 1946), as well as an admirable exposition of the static theory of value. Perhaps the most important of modern works is Paul A. Samuelson's *Foundations of Economic Analysis* (Harvard University Press, 1947), though it is not accessible to those without a fair mathematical equipment. It contains a definitive account of the theory of maximization, and the first systematic exposition of dynamic theory using the tool of difference equations. Baumol's *Economic Dynamics* (Macmillan, 1951) is a very useful introduction to this field, as is R. F. Harrod's *Toward a Dynamic Economics* (Macmillan, 1948).

The field of welfare economics has shown great activity in recent years. Pigou's *Economics of Welfare* (4th ed., Macmillan, 1932), though now quite out of date, is still a "classic" work in the field. Students interested in modern welfare economics should read works by Reder,[35] Myint,[36] Little,[37] and Arrow.[38] Arrow's work in particular represents a somewhat new departure in economics, using the apparatus of mathematical logic.

The nearer we get to the present day the more important becomes periodical literature as a source of important contributions. The *Quarterly Journal of Economics,* published in Cambridge, Mass., founded in 1885, is probably the oldest periodical specializing in economic theory. The *Economic Journal,* organ of the Royal Economic Society and published in London, was founded in 1891. The *Journal of Political Economy* (founded in 1892 and published at the University of Chicago) and the *American Economic Review* (organ of the American Economic Association, founded in 1911) often contain important theoretical articles, but tend predominantly in the direction of applied studies. *Economica* (founded in 1921) is the organ of the London School of Economics. The *Review of Economic Studies* (founded in 1933) is a frequent vehicle for the younger theorists and contains many important articles. *Econometrica* (organ of the Econometric Society and founded in 1933) is the principal vehicle for articles in mathematical and statistical economics. The *Review of Economic Statistics* (Harvard University) is also important in this field. There are several "local" journals which from time to time have articles of general interest: the *Canadian Journal of Economics and Political*

[35] Melvin W. Reder, *Studies in the Theory of Welfare Economics,* Columbia University Press, 1947.

[36] Hla Myint, *Theories of Welfare Economics,* Harvard University Press, 1948.

[37] I. M. D. Little, *A Critique of Welfare Economics,* New York, Oxford University Press, 1950.

[38] Kenneth Arrow, *Social Choice and Individual Values,* New York, Wiley, 2nd ed., 1963.

Science, the *Economic Record* (Australia), the *South African Journal of Economics, The Manchester School,* are especially to be recommended. The *Zeitschrift für Nationaloekonomie* (Vienna) and the *Giornale degli economisti* (Italy) are also of importance. *Kyklos* (Zurich) is an international journal with articles in several languages. It would be impossible to attempt a bibliography of the important articles in these journals in the small space of this appendix; the student whose interests lie in any special field will soon have to make such a bibliography for himself. The book reviews in these journals (especially in the *American Economic Review*) in themselves form a most valuable bibliography both for general theory and for all the special fields. The American Economic Association now publishes *The Journal of Economic Abstracts,* invaluable in the light of the increasing volume of periodical literature.

In recent years the publications of specialized economic research agencies, such as the Cowles Foundation at Yale University and the National Bureau of Economic Research in New York, and publications of government departments, such as the U.S. Department of Commerce, have become increasingly important, especially on the empirical side of economics.

The volumes of collected articles published by the American Economic Association are a useful introduction to the periodical literature. The *Survey of Contemporary Economics,* Vols. I and II, also published by the American Economic Association, is a useful guide to recent economic thought and bibliography in many fields.

The list might be continued almost indefinitely, and it is not the purpose of this appendix to provide the student with a complete bibliography. It is hoped, however, that the works mentioned above may provide an essential point of departure for further inquiry.

INDEX